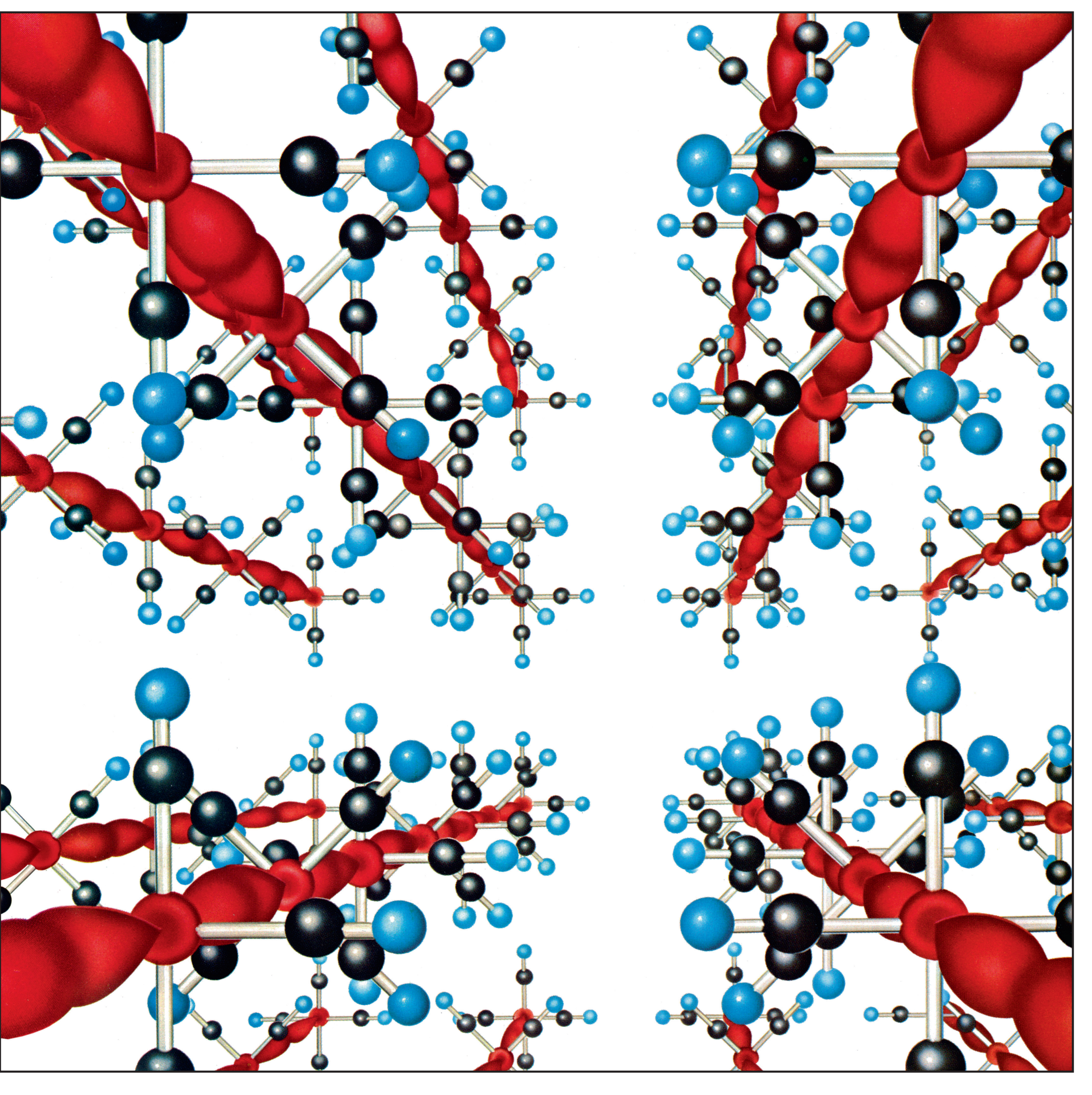
INVESTIGACION Y CIENCIA

Edición en español de **SCIENTIFIC AMERICAN**



CONDUCTORES DE CADENA LINEAL

Diciembre 1979

200 PTAS.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 39
Diciembre 1979

INVESTIGACION Y
CIENCIA
Edición en español de SCIENTIFIC AMERICAN

SCIENTIFIC AMERICAN

COMITE DE REDACCION Gerard Piel (Presidente), Dennis Flanagan, Francis Bello, Philip Morrison, Judith Friedman, Brian P. Hayes, Paul W. Hoffman, Jonathan B. Piel, John Purcell, James T. Rogers, Armand Schwab, Jr., Jonathan B. Tucker y Joseph Wisnovsky

DIRECCION EDITORIAL Dennis Flanagan
DIRECCION ARTISTICA Samuel L. Howard
PRODUCCION Richard Sasso
DIRECTOR GENERAL George S. Conn

INVESTIGACION Y CIENCIA

DIRECTOR Francisco Gracia Guillén
REDACCION José María Valderas Gallardo (Redactor Jefe)
Encarna de Blas (Secretaria de Redacción)
PRODUCCION César Redondo Zayas
PROMOCION Y PUBLICIDAD Elena Sánchez-Fabrés
PROMOCION EXTERIOR Pedro Clotas Cierco

EDITA Prensa Científica, S.A.
Calabria, 235-239
Barcelona-29 (ESPAÑA)

Colaboradores de este número:

Asesoramiento y traducción:

Amando García: *Conductores de cadena lineal;* José M.ª Vidal: *El microscopio acústico;* Miguel Viñas: *Legionelosis;* Carlos Oppenheimer: *La membrana fotosintética;* Montserrat Domingo: *Los lagos de lava del Kilauea;* Joan Domènec Ros: *Evolución de los Crocodilios;* Luis Bou: *Matemática constructivista; Juegos matemáticos;* E. Myro y Alicia Menéndez: *Taller y laboratorio;* Manuel Puigcerver: *Lluvia ácida.*

Ciencia y sociedad:
Francisco J. Yndurain

Libros:
Antonio Prevosti
y Mercedes Durfort



LA PORTADA

Vista interna de un material cristalino que es un buen conductor de la electricidad únicamente en una determinada dirección [véase "Conductores de cadena lineal", en este mismo número]. El material en cuestión es una sal de tetracianuro de platino. Las largas estructuras en forma de columna que parecen alejarse del plano frontal de la figura son otras tantas pilas de moléculas de tetracianuro de platino. Los átomos de carbono (*gris*) y los átomos de nitrógeno (*azul*) se han representado esquemáticamente en forma de esferas, en tanto que los átomos de platino (*rojo*), que están situados en el centro de cada molécula, vienen representados por los orbitales que definen la distribución de electrones en dicho átomo. Debido al hecho de que estos orbitales se extienden a uno y otro lado del plano de la molécula, los orbitales de las moléculas adyacentes en una misma columna se solapan entre sí. Este solapamiento da lugar a que la conductividad del cristal en una dirección paralela a dichas columnas sea relativamente elevada; la conductividad en una dirección perpendicular a la anterior resulta ser unas 100.000 veces inferior.

Suscripciones:
Prensa Científica, S.A.
Calabria 235-239
Barcelona-29 (España)
Teléfono 322 05 51 ext. 41

Condiciones de suscripción:
España:
Un año (12 números): 2200 pesetas
Extranjero:
Un año (12 números): 36 U.S.$
Ejemplar atrasado ordinario:
225 pesetas
Ejemplar atrasado extraordinario:
340 pesetas

Distribución para España:
Distribuciones de Enlace, S.A.
Ausias March, 49, Barcelona-10

Distribución para los restantes países:
Editorial Labor, S.A.
Calabria, 235-239, Barcelona-29

Publicidad:
Madrid:
Gustavo Martínez Ovin
Avda. de Moratalaz, 137, Madrid-30
Tel. 430 84 81
Cataluña:
Miguel Munill
Balmes, 191, 2.º, 2.ª, Barcelona-6
Tels. 218 44 45 y 218 40 86

Controlado
por O.J.D.



PROCEDENCIA DE LAS ILUSTRACIONES

Ilustración de la portada de George V. Kelvin

Página	Fuente
8-14	Roberto Daniel Lozano y A. Sellés
17	George V. Kelvin
18-19	Xerox Corporation
20-28	George V. Kelvin
31	Daniel Rugar, Universidad de Stanford
32-35	Dan Todd
36	Abdullah Atalar Universidad de Stanford
37-38	Calvin F. Quate, Universidad de Stanford
49	Joseph E. McDade, Center for Disease Control, U.S. Public Health Service (*arriba, izquierda*); William B. Cherry, Center for Disease Control (*arriba, derecha*); Francis W. Chandler, Center for Disease Control (*abajo*)
50	Francis W. Chandler
51-58	Bunji Tagawa
63	Kenneth R. Miller, Universidad de Harvard
64-65	Jerome Kuhl
68	Nam-Hai Chua, Universidad Rockefeller
69	Kenneth R. Miller, Universidad de Harvard (*arriba*); Jerome Kuhl (*abajo*)
70	Jerome Kuhl
71-72	Kenneth R. Miller
73	Kenneth R. Miller (*arriba y abajo izquierda*); L. Andrew Staehelin, Universidad de Colorado (*abajo, derecha*)
77	Dallas L. Peck y Thomas L. Wright, U.S. Geological Survey; Robert W. Decker, Hawaiian Volcano Observatory
78-79	Dallas L. Peck y Thomas L. Wright, U.S. Geological Survey; Robert W. Decker, Hawaiian Volcano Observatory (*izquierda*) Walken Graphics (*derecha*)
80-81	Walken Graphics
82	Dallas L. Peck y Thomas L. Wright, U.S. Geological Survey; Robert W. Decker, Hawaiian Volcano Observatory
83-84	Walken Graphics
89	Eric Buffetaut, CNRS
90-98	Patricia J. Wynne
101	Scala / Editorial Photocolor Archives
102-114	Ilil Arbel
117	James Young
118-121	Michael Goodman

ISSN 0210-136X
Dep. legal: B. 38.999-76
Fotocomposición Tecfa
Guizpúzcoa, 36 (local 1) Barcelona-20
Cayfosa. Santa Perpetua de Moguda
Barcelona
Printed in Spain - Impreso en España

Los autores

ROBERTO DANIEL LOZANO ("Diferencias de color") es jefe de la División Optica del Departamento de Física del Instituto Nacional de Tecnología Industrial de Buenos Aires, Argentina. Aunque es licenciado en matemáticas aplicadas, su campo de trabajo ha estado siempre relacionado con la óptica y de ella en particular con la luz y el color. Fue becario del Laboratorio Nacional de Física (NPL) de Gran Bretaña y del Instituto Federal de Física (PTB) de Alemania. Es miembro consultivo de distintos comités técnicos de la Comisión Internacional del Alumbrado (CIE). Sus trabajos en el tema han sido diversos: visión de color (estudio de la aditividad y efectos de adaptación de las funciones de igualación de color para un campo visual extenso), apariencia (blancura y brillo), diferencias de color, materiales retrorreflectores, instrumentación y alimentos, entre otros. La revista reseñó recientemente su libro sobre *El color y su medición*.

ARTHUR J. EPSTEIN y JOEL S. MILLER ("Conductores de cadena lineal") se han dedicado a estudiar los materiales moleculares conductores durante casi una década. Epstein es un físico del Centro de Investigación Webster de la Xerox Corporation en Rochester, Nueva York. Se licenció en el Instituto Politécnico de Brooklyn. En el año 1971 obtuvo el grado de doctor en físicas en la Universidad de Pennsylvania. Trabajó en la Mitre Corporation, incorporándose al personal de investigación de la Xerox en 1972. Miller es un químico que trabaja actualmente en la Occidental Research Corporation de Irvine, California. Tras estudiar su licenciatura en la Universidad del estado de Wayne, en 1971 recibió su título de doctor en química inorgánica en la Universidad de California en Los Angeles. Posteriormente trabajó en la Universidad de Stanford y más tarde se unió a la Xerox con el fin de estudiar la química de los materiales moleculares de cadena lineal. Finalmente, en el curso de este año, ingresó en la Occidental Research.

CALVIN F. QUATE ("El microscopio acústico") enseña ingeniería eléctrica y física aplicada. Dirige el Departamento de Física Aplicada en la Universidad de Stanford. Se graduó en ingeniería eléctrica en la Universidad de Utah. Obtuvo su Doctorado en Física en 1950 en Stanford. En 1949 pasó a formar parte del equipo de investigación técnica de los Laboratorios Bell, en donde llegó a ser director asociado de investigación electrónica. Después de estar 10 años en los Laboratorios Bell pasó a Albuquerque, N. M., a la Sandia Corporation, de donde fue vicepresidente y, más tarde, director de investigación. Se trasladó a Stanford en 1961.

DAVID W. FRASER y JOSEPH E. McDADE ("Legionelosis") pertenecen a la plantilla del Centro para el Control de las Enfermedades (CDC) del Servicio de Salud Pública de los Estados Unidos en Atlanta. Fraser es jefe de la rama especial de patógenos de la división de enfermedades bacterianas en el Centro de Epidemiología. Se graduó en el Haverford College y se doctoró en la Facultad de Medicina de Harvard en 1969. Tras una breve estancia en la Universidad de Pennsylvania pasó a la sección de proyectos especiales como epidemiólogo en la Mayo Graduate School of Medicine, en Rochester. Volvió a la Universidad de Pennsylvania como jefe residente al par que docente en la Facultad de Medicina. Se trasladó al CDC en 1977. McDade es investigador microbiólogo en la rama de lepra-rickettsias de la división de virología. Una vez obtenido su doctorado en microbiología por la Universidad de Delaware en 1967, se dedicó a lo largo de dos años al estudio de las infecciones por rickettsias en el Centro Biológico del Ejército de los Estados Unidos en Fort Detrick. Trabajó otros dos años como director del departamento de producción celular de "Microbiological Associates Inc". De 1971 a 1975 perteneció adscrito a la Facultad de Medicina de la Universidad de Maryland, pasando la mayoría de ese tiempo en el Cairo y Etiopía estudiando los ciclos extrahumanos del tifus epidémico. Ingresó en el CDC en 1975.

KENNETH R. MILLER ("La membrana fotosintética") es profesor adjunto de biología en la Universidad de Harvard. Sus trabajos de pregraduado los realizó en la Universidad Brown, doctorándose en biología por la Universidad de Colorado en Boulder en 1974. Su investigación se dirige al estudio de la función de la membrana fotosintética en la reacción luminosa de la fotosíntesis, técnicas experimentales de alta resolución en microscopía electrónica y la bioquímica de las macromoléculas asociadas a la membrana.

DALLAS L. PECK, THOMAS L. WRIGHT y ROBERT W. DECKER ("Los lagos de lava del Kilauea") comenzaron a estudiar los lagos de lava fundida en 1963, cuando se hallaban en el Hawaiian Volcano Observatory. Peck es geólogo jefe del U. S. Geological Survey. Cursó sus estudios en el Instituto de Tecnología de California y se doctoró por la Universidad de Harvard en 1960. Realizó luego trabajos para el Geological Survey en Oregón occidental, en la zona de California de granito batolítico de Sierra Nevada y en Hawai. Wright es geólogo contratado por el Geological Survey. Estudió en el Pomona College y se doctoró por la Universidad John Hopkins en 1961. Prestó después servicios para el Geological Survey en Cascade Range de Washington y en Hawai. Desde que salió de Hawai, en 1969, ha venido estudiando basaltos de la luna, de la cresta del Atlántico y del río Columbia. Decker es jefe de investigación del Hawaiian Volcano Observatory. Estudió en el Instituto de Tecnología de Massachusetts y se doctoró por la Escuela Superior de Minas de Colorado en 1953. Tras ser profesor e investigador de volcanología en Indonesia se incorporó al Dartmouth College. A principios del año en curso recibió el nombramiento de director del Volcano Observatory.

ERIC BUFFETAUT ("Evolución de los Crocodilios") forma parte del equipo de investigación del Centre National de la Recherche Scientifique, de Francia, y trabaja en el Laboratorio de Paleontología de los Vertebrados de la Universidad de París. Nacido en Normandía, se interesó por los fósiles desde niño. Estudió geología y paleontología en la Universidad de París y se doctoró en 1974 con una tesis sobre los Crocodilios fósiles del Cretácico superior del Níger. Las investigaciones de Buffetaut sobre la historia evolutiva y biogeográfica de los Crocodilios le ha llevado a realizar trabajo de campo en Francia, Canadá, Tunicia, Malí y Paquistán.

ALLAN CALDER ("Matemática constructivista") es lector de matemáticas en el Birbeck College de la Universidad de Londres. Formado en esta misma universidad, se graduó en matemática en 1968, y obtuvo el doctorado por una tesis en topología, en 1970. Desde esa fecha ha profesado en diversas universidades, entre ellas, la Universidad de Carleton, de Ottawa, en la estatal de Louisiana, la Universidad de Essex y la de Missouri. Actualmente es profesor visitante asociado de la Universidad estatal de Nuevo México.

Hace...

José M.ª López Piñero

...cuatrocientos años

Nació en Jaén Juan Gutiérrez de Godoy, que llegaría a ser una de las figuras médicas más destacadas en la España de la primera mitad del siglo XVII. Doctorado en la Universidad de Alcalá, ejerció sucesivamente en Alcalá la Real, Jaén y Córdoba. En 1645, su carrera profesional culminó con el nombramiento de médico de cámara de Felipe IV, cargo que ocupó hasta su muerte once años más tarde.

La principal obra de Gutiérrez de Godoy es el libro titulado *Tres discursos para probar que están obligadas a criar a sus hijos a sus pechos todas las madres* (1629), generalmente considerado como uno de los mejores textos sobre puericultura de la época. Frente a la costumbre entonces vigente entre las mujeres nobles, o simplemente acomodadas, de entregar sus hijos a nodrizas mercenarias, se trata de una enérgica defensa del derecho del niño a la lactancia materna, "quando las madres tienen buena salud, fuerças y buen temperamento, buena leche y suficiente para alimentar (a sus hijos)". En el primer "discurso" expone argumentos de carácter biológico, basados en las concepciones fisiológicas de la medicina galénica tradicional; de acuerdo con ellas, intenta demostrar que la leche de madre es la mejor y más conveniente para el niño porque es de la misma "naturaleza" que el alimento que ha recibido en el útero. En los otros "discursos" ofrece razones higiénicas, sociales y morales apoyadas en la observación de las condiciones reales de la lactancia mercenaria, que describe minuciosamente.

Gutiérrez de Godoy publicó, además, unos comentarios sobre el texto aristotélico. *De memoria et reminiscentia* ("Sobre la memoria y el recuerdo") y varias obras menores de contenido clínico, escritas en respuesta a diferentes consultas de otros médicos.

...trescientos años

La epidemia de peste bubónica que padeció la ciudad de Málaga durante 1678 y 1679 motivó la publicación de una serie de estudios que reflejan las ideas epidemiológicas de la época, así como la práctica y la organización sanitarias entonces vigentes.

El primero de dichos estudios fue el *Tratado de la peste de Málaga* (1679) de Bernardo Francisco de Acevedo, médico residente en la ciudad que describe en estos términos el comienzo de la epidemia: "Málaga gozaba el año de 1678 de la salud más perfecta, aunque con la zozobra de la peste que se padecía en Orán desde el año antecedente. Sin embargo de haber puesto todo el cuidado y vigilancia que exigían tales circunstancias para evitar el peligro que le amenazaba por su proximidad al mal, no fueron bastantes a impedir que aportase a su puerto el día 28 de mayo de 78 una saetía que, ocultando el rumbo que traía desde Orán, se admitió al comercio, conduciendo con los géneros la dañina peste que temían. Saltaron en tierra algunos marineros... Al segundo día murió con mucha celeridad un muchacho de catorce años de la misma embarcación. Al otro día enfermó un marinero de la referida embarcación y murió a la entrada del quinto día con calentura maligna, carbunclos en el cuerpo y una seca, landre o bubón en una ingle; dos hermanas tuvieron después la misma calentura y síntomas, un carbunclo cada una en la pierna y una landre en la correspondiente ingle...".

Acevedo y casi todos los demás médicos de la ciudad consideraron que se trataba de una "pestilencia contagiosa", por lo que se organizó un "hospital de separación" o de cuarentena. Sin embargo, la comisión médica enviada por la Real Chancillería de Granada fue de la opinión contraria. La encabezaba el catedrático granadino Marco Antonio de Checa, que publicó una *Carta apologética, en que se prueba que la enfermedad que corrió este año pasado en la ciudad de Málaga, no fué peste* (1679). En consecuencia, se desmontó el hospital cuarentenario y todas las demás medidas preventivas.

Recrudecida la epidemia de forma dramática, el médico malagueño Pedro Biosca Casanova criticó el dictamen de los comisionados de la Chancillería y reiteró el diagnóstico de Acevedo en una *Carta antiapologética: respuesta a otra del Doctor Marco Antonio de Checa... en que se defiende y prueba haber sido peste la enfermedad que corrió este año pasado de 78 en la ciudad de Málaga* (1679). Ante la situación, intervino el poder real a través del Protomedicato, que envió a Diego Blanco Salgado, una autoridad en el tema. Este confirmó la opinión de los médicos malagueños y publicó un brillante informe bajo el título de *Tratado de la epidemia pestilente que padeció la ciudad de Málaga el año de 1678 y 1679*.

...doscientos años

Nació en Granada José Mariano Vallejo, uno de los escasos matemáticos españoles de relieve durante las primeras décadas del siglo XIX.

Vallejo estudió en la Universidad de Granada, aunque su formación matemática fue fundamentalmente autodidacta. En 1801 fue profesor sustituto de geometría práctica en la Real Academia de San Fernando y colaboró en diversos trabajos técnicos por encargo de dicha institución. Al año siguiente ganó por oposición la cátedra de matemática y fortificación del Real Seminario de Nobles de Madrid, puesto desde el que realizó la mayor parte de su obra científica. En la Guerra de Independencia trabajó como técnico al servicio del Cuerpo de Artillería. Al terminar el trienio liberal, en 1823, se exilió, como tantos otros científicos españoles. Recorrió varios países europeos y acabó asentándose en París, donde enseñó matemática y mantuvo una excelente amistad con Pierre Simon Laplace, que se interesó por sus contribuciones. En 1832, poco antes de la muerte de Fernando VII, regresó a España. Durante las dos décadas que le quedaban de vida se esforzó en divulgar las nuevas corrientes pedagógicas europeas. Llegó a inspector de instrucción pública y promovió la creación de escuelas normales. Intervino en la fundación de la Academia de Ciencias y en la del Ateneo de Madrid, de cuya sección de matemática y física fue presidente.

Vallejo publicó en primer término unas *Adiciones a la Geometría de D. Benito Bails* (1806) en las que, sin estar todavía familiarizado con la obra de los más destacados matemáticos del momento, desarrolló algunos aspectos en una línea semejante a la de éstos.

Santiago Garma, el mejor estudioso de la obra de Vallejo, considera que su publicación más importante es la *Memoria sobre la curvatura de las líneas* (1807). Destaca no solamente el interés de su contenido, sino su carácter de monografía original en medio de un pano-

rama bibliográfico dominado por manuales y compendios. Está dedicada principalmente a las evolutas y concluye con la siguiente afirmación: "De toda curva algebraica se puede obtener la expresión de su radio de curvatura, y las evolutas de estas curvas serán todas rectificables, es decir, se podrá señalar la longitud de sus curvas por una fórmula algebraica".

Vallejo compuso dos obras de síntesis. Su *Tratado elemental de Matemáticas* (1813) consta de cinco volúmenes en los que se expone desde la aritmética hasta el cálculo diferencial e integral. Fue reeditado en dos ocasiones, la última de ellas en 1821 bajo la égida del gobierno liberal. El *Compendio de Matemáticas* (1819) es un texto mucho más breve en el que Vallejo intentó recopilar toda su producción anterior. Concedía gran importancia al proceso del razonamiento y al método de estudio en matemática, por lo que en ambas obras añadió un capítulo de lógica en el que explica el lenguaje que va a utilizar.

...cien años

Nació en Barcelona August Pi i Sunyer. Hijo de un catedrático de la Facultad de Medicina de Barcelona, estudió en dicho centro, licenciándose en 1899. Desde muy joven frecuentó el Laboratorio Municipal de Bacteriología de la capital catalana, donde Ramón Turró fue su auténtico maestro. En 1904 obtuvo por oposición la cátedra de fisiología de la Facultad de Medicina de Sevilla, pero permaneció en Barcelona con diversas comisiones. En 1916 sucedió a Coll i Pujol en la cátedra de fisiología barcelonesa y, cuatro años después, se fundó el Instituto de Fisiología del que fue nombrado director. El notable equipo que encabezó en estas instituciones quedó deshecho por la guerra civil. Pi i Sunyer se exilió y, tras una breve estancia en París, se asentó en Caracas como profesor de fisiología y, más tarde, de bioquímica. Allí fundó el Instituto de Medicina Experimental, núcleo originario de la moderna escuela venezolana de investigación biomédica. El fisiólogo catalán falleció en la ciudad de México en 1965.

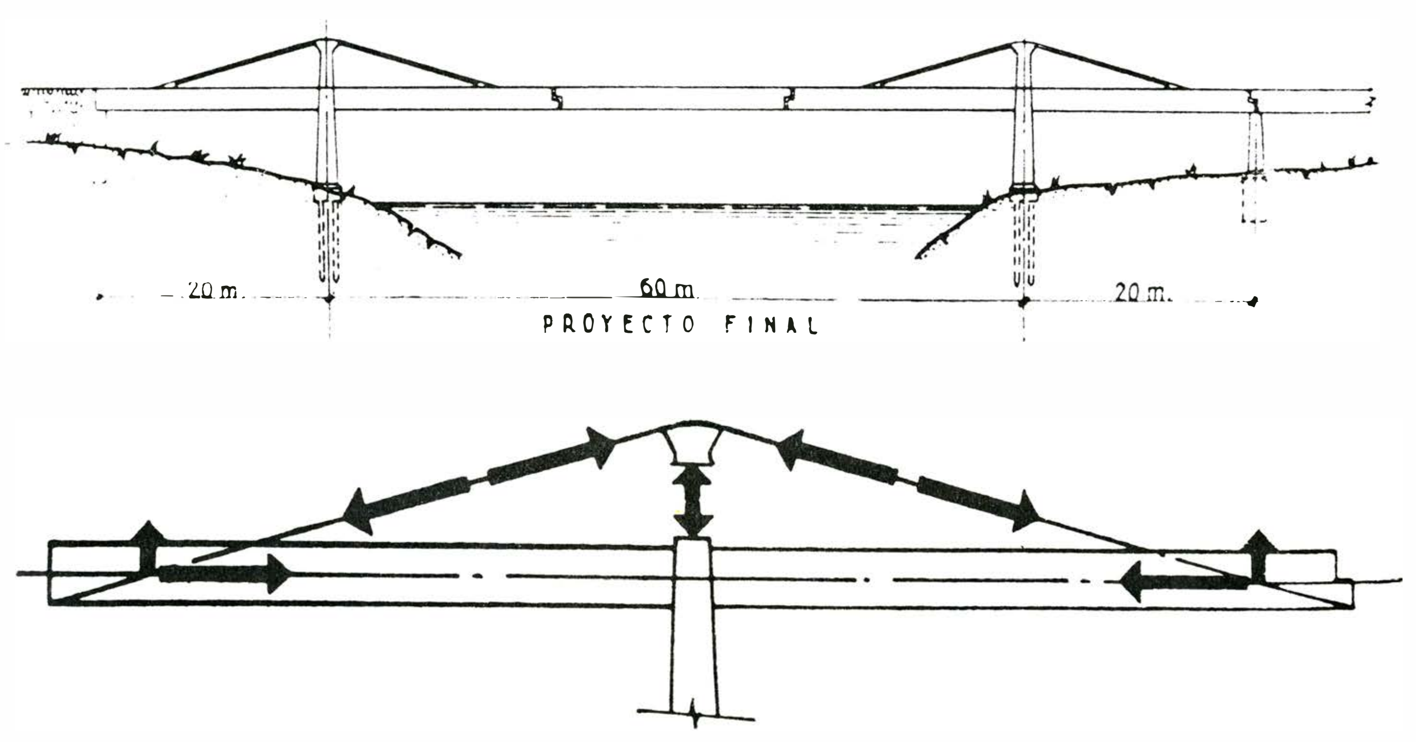
Dejando aparte su tesis doctoral que versó sobre la vida anaerobia (1901), los primeros trabajos de Pi i Sunyer estuvieron encuadrados en los intentos de Turró de formular una teoría explicativa de los mecanismos de inmunidad natural basada en la identidad de dichos fenómenos y los digestivos. Aunque los resultados no fueron satisfactorios, sirvieron de puntos de partida de sus posteriores investigaciones sobre los reflejos nerviosos de adaptación, terreno en el que realizó sus más destacadas aportaciones. A partir de 1918 estudió el mecanismo reflejo regulador de los movimientos respiratorios y pudo demostrar el papel que desempeñaban los quimiorreceptores periféricos con aferencia vagal. También investigó las condiciones nerviosas vegetativas que intervienen en la regulación de la glucemia (1922).

Realizó alguna otra contribución de cierto relieve, como la demostración de la acción inhibidora que la sangre urémica tiene sobre la secreción urinaria (1917). No obstante, la importancia de su obra reside más bien en la amplia serie de trabajos de síntesis y de manuales que publicó, basados en una excelente información. Recordaremos, entre los más tempranos, *Los mecanismos de correlación fisiológica* (1920) y, entre los últimos, el tratado de *Fisiología humana* (1962) en colaboración con Jaume Pi i Sunyer. Convencido catalanista, Augusto Pi i Sunyer fue fundador de la Societat de Biologia (1912) y director de sus *Treballs,* participó activamente en los congresos de médicos de lengua catalana y contribuyó de forma muy destacada al esfuerzo destinado a convertir esta última en un medio normal de comunicación científica.

...cincuenta años

Eduardo Torroja Miret dio un curso sobre "Hormigón Armado" en la Escuela de Arquitectura de Madrid y presentó, en el Congreso Internacional de Puentes celebrado en Viena, la comunicación titulada "L'emploi des cábles d'acier dans les constructions en béton armé".

Nacido en Madrid el año 1899 y muerto en la misma ciudad en 1961, Torroja Miret era hijo del gran matemático Eduardo Torroja Caballé. Estudió en la Escuela de Ingenieros de Caminos, ter-



Acueducto de Tempul, sobre el río Guadalete, en cuya construcción (1926) empleó ya Eduardo Torroja Miret la idea del pretensado. Aparte del proyecto inicial y del final, se reproduce un esquema del sistema de pretensado de su estructura.

minando la carrera en 1923. Desde su primera época profesional manifestó su inquietud por la búsqueda de nuevas formas, desde el punto de vista matemático y desde el estético, así como de nuevas bases tecnológicas. Ello le había de conducir a convertirse en uno de los adelantados mundiales en el campo del hormigón pretensado. Ya en 1926, empleó la idea del pretensado en la construcción del acueducto de Tempul, sobre el río Guadalete, parte del abastecimiento de aguas de Jerez de la Frontera. Para evitar la fisuración y limitar el alargamiento de los tirantes, colocó en la parte alta de las pilas unos dispositivos de elevación que permitían tensar los cables. En la misma línea, otra de sus obras tempranas más destacadas fue el proyecto y construcción, en 1933, de la cubierta del mercado de Algeciras, cuya forma totalmente nueva fue posible gracias a una original aplicación de la técnica del pretensado; se trata de un casquete esférico de hormigón armado apoyado sobre ocho pilares, con un diámetro libre de unos cuarenta y ocho metros de luz y un espesor de hormigón que solamente llega a los nueve centímetros. Sin embargo, la primera obra en la que Torroja aplicó con plena conciencia la técnica del pretensado –que Freyssinet había expuesto en 1935– fue el acueducto de Alloz (1939), después seguido por una amplísima serie de proyectos y construcciones.

Preocupado por fomentar en España los progresos de la construcción y sus materiales, Torroja fundó en 1934 el Instituto de la Construcción y Edificación, centro privado que reunió un grupo de ingenieros, arquitectos y constructores. Ese mismo año comenzó a publicarse la revista *Hormigón y Acero*. Tras la guerra civil, el Instituto fue integrado en el nuevo Consejo Superior de Investigaciones Científicas, fusionándose después con el Instituto del Cemento, bajo la dirección del propio Torroja. La adhesión de éste al régimen político de Franco facilitó también su incorporación en 1939 al profesorado de la Escuela de Ingenieros de Caminos, su nombramiento en 1940 como director del Laboratorio Central de Ensayo de Materiales de Construcción y su ingreso en la Real Academia de Ciencias (1944).

Autor de más de un centenar de trabajos, entre los publicados en España y los aparecidos en revistas extranjeras, Torroja alcanzó un gran prestigio internacional. Dio cursos y conferencias en la mayor parte de los países europeos, Estados Unidos y Latinoamérica y fue nombrado doctor "honoris causa" o miembro honorario de importantes instituciones técnicas de todo el mundo.

Diferencias de color

La tarea de transformar la medición subjetiva del color en términos objetivos requiere una laboriosa interpretación del mecanismo visual y psicológico que transforma los estímulos del medio en sensaciones

Roberto Daniel Lozano

La realidad del mundo visual que nos rodea está llena de color. La mayoría de los seres humanos están tan habituados a ello que es común la aceptación de este fenómeno como algo natural, parte de la realidad misma. Si a un observador desprevenido (con visión del color normal) se le pide que diga qué color tiene un objeto, rápidamente responderá, según su experiencia, que es rojo, verde, azul, etcétera. Inmediatamente atribuirá a ese objeto el color indicado como parte de sus propiedades intrínsecas. Así dirá, por ejemplo, la manzana *es* roja, la planta *es* verde, la tela *es* amarilla.

Supóngase que nos encontramos frente al último caso: el de una tela amarilla. ¿Es realmente amarilla? y si lo es, ¿todos los amarillos son iguales? En principio la respuesta es sencilla: la experiencia nos dice que no. En realidad, podríamos afirmar, sin pensar mucho, que existen infinidad de diferentes colores aun dentro de un mismo tono. Puede haber amarillos más rojizos o más verdosos, claros u oscuros y aún más o menos amarillos.

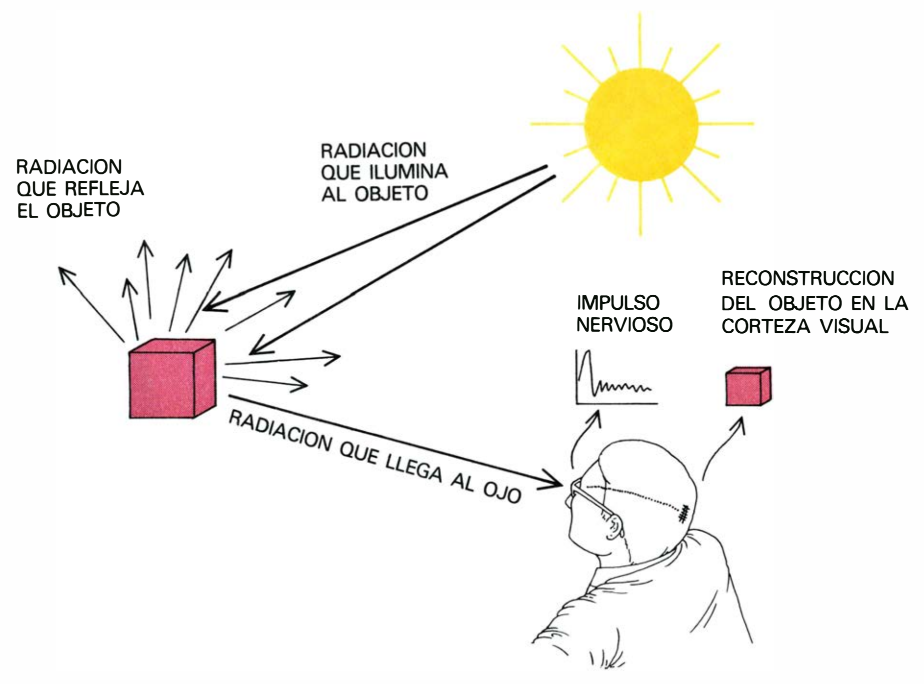
Este es un problema que preocupa mucho a las industrias que requieren de su producto una uniformidad más o menos estricta. Por ejemplo, es el caso de la textil, ¿Qué comprador aceptaría una tela cuyo color no fuera uniforme, salvo naturalmente los casos en que este efecto esté realizado intencionadamente? O en la producción automotriz, ¿quién compraría un auto en el cual un guardabarro tuviera un color similar pero no exactamente igual al del resto de la carrocería?

Este tipo de problemas ha llevado a los investigadores al estudio de la ciencia del color y su medición –la colorimetría–. Parte de esta ciencia está dedicada al problema de la medición de diferencias de color. Para entender la naturaleza de este asunto es necesario describir cómo se mide un color para luego pasar a la medición de las diferencias. El color es un atributo sensorial del hombre y, en particular, de la visión. Lo que un observador ve es el producto de la luz que sale de un emisor, se refleja en el (o transmite a través del) objeto y llega a la retina del ojo transformándose en impulsos nerviosos que se transmiten a la corteza cerebral para dar lugar a la sensación.

Desde un punto de vista físico, el sol no es más que un fuente radiante, un objeto que emite radiación electromagnética cuya característica fundamental es la composición espectral y la distribución de intensidades a lo largo del espectro. El ojo, por otra parte, sólo sería un detector con una respuesta espectral muy especial, que nos permite "ver" –esto es: detectar– sólo una muy pequeña porción del espectro electromagnético. (Entre 380 y 770 nanómetros, aproximadamente, de longitud de onda.) Otras radiaciones de distintas longitudes de onda no pueden ser vistas por los observadores humanos. Por ejemplo, "sentimos" la radiación infrarroja, cuando ella es absorbida por la piel a través de nuestros sensores térmicos cutáneos, pero no podemos verla.

De lo dicho se deduce que, para identificar fisicamente una radiación determinada, basta dar la curva de distribución espectral de sus intensidades. Distintos tipos de fuentes producen diferentes espectros. A esta distribución la simbolizaremos $E(\lambda)$. Vayamos ahora a los objetos. Estos pueden absorber, transmitir o reflejar la luz. Si son opacos no podrán transmitirla. La luz, y llamamos luz a la radiación electromagnética visible, podrá a su vez ser o no difundida.

La difusión de la luz en un medio es el fenómeno de desvío del camino óp-



LA RADIACION ELECTROMAGNETICA proveniente de la fuente (en este caso: el sol) llega al objeto donde se refleja. La radiación reflejada se propaga entonces hasta el ojo humano donde incide en la retina; una vez allí dará lugar al nacimiento de un impulso eléctrico que viaja a través de las conexiones nerviosas hasta la corteza, originando la percepción de la imagen cromática del objeto.

tico, por interacción con materia interpuesta. No es posible describir aquí la forma en que esto ocurre, pero dígase que ello depende de la naturaleza misma de que está compuesto el objeto y que también es dependiente de la longitud de onda de la radiación incidente. A los efectos del presente trabajo simplificaremos diciendo que si el objeto es opaco, o sea, no pasa luz a través del mismo, la luz puede ser reflejada o absorbida. En casos extremos se tiene un espejo perfecto, donde toda luz es reflejada en la dirección especular, a saber, con igual ángulo respecto de la normal a la superficie que el de incidencia. Si la superficie difiere del espejo perfecto habrá una componente reflejada difusa que irá creciendo cuanto más difusor ("mate" en la expresión vulgar) sea la misma.

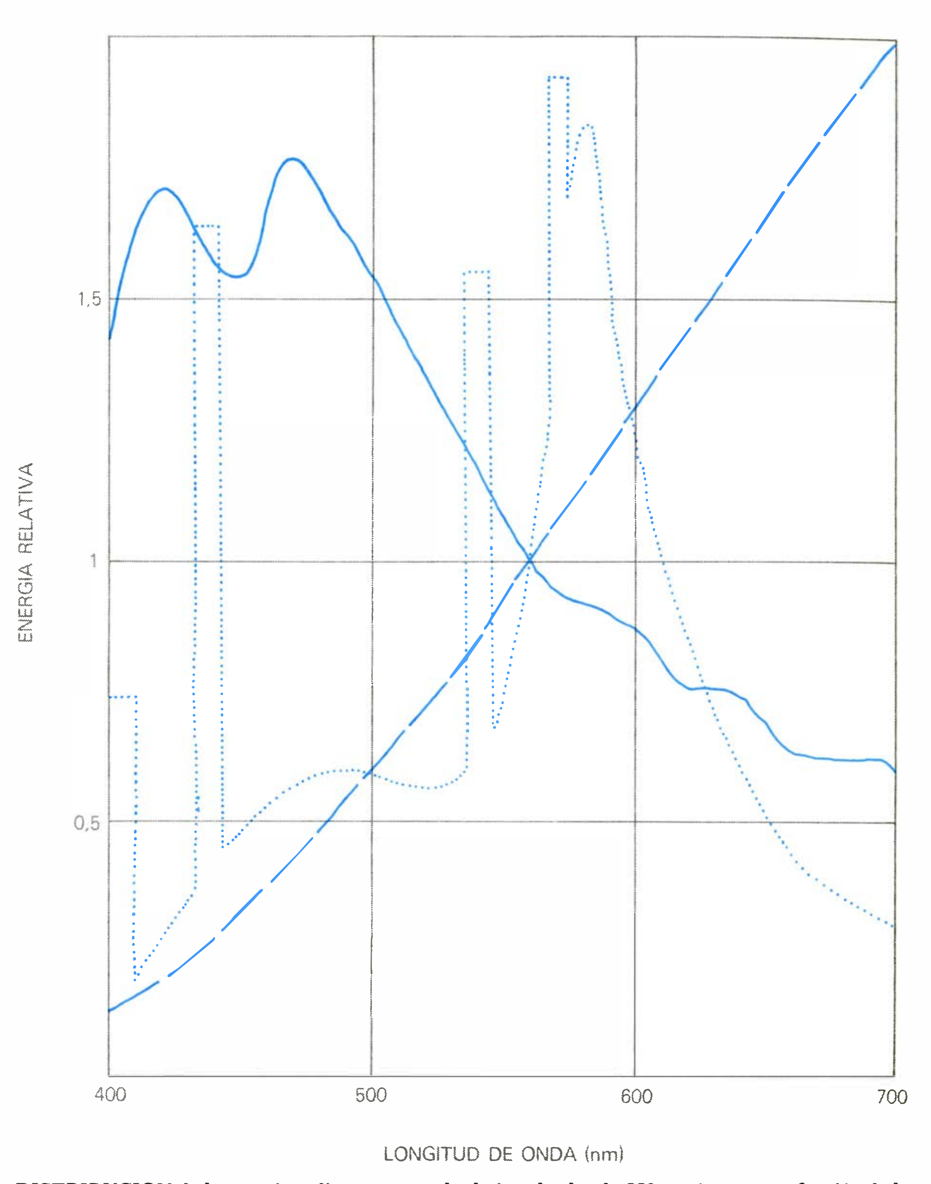
Es importante notar que la reflexión difusa se produce por la penetración de la luz en el medio (objeto). Al interactuar con éste, se produce absorción de algunas radiaciones además de la desviación del haz incidente (dependiendo este efecto de la longitud de onda).

La reflexión difusa es, por consiguiente, la que produce la sensación de color de los objetos. La reflexión especular nos da la impresión de *brillo*, que sólo depende de la característica de la superficie externa del objeto (si es pulida o no) y del índice de refracción. (Para entender el concepto de brillo aquí descrito asóciese esa propiedad al efecto que uno observa al contemplar una pintura brillante y otra mate del mismo color.)

La variación del flujo radiante reflejado difusamente por el objeto de la longitud de onda recibe el nombre de reflectancia espectral de un objeto y se nota como $R(\lambda)$. El observador ve el producto de la distribución espectral de la luz $E(\lambda)$ que ilumina el objeto por la reflectancia espectral del objeto observado $R(\lambda)$. El resultado de esta interacción física se simboliza por $P(\lambda)$.

El ojo humano tiene la propiedad de poder igualar cualquier color con la suma de tres colores primarios. (Se entiende por colores primarios aquellos que son independientes entre sí, esto es, que cualquier suma de dos de ellos no pueden igualar al tercero.) Si se suman luces será una mezcla aditiva; si se pinta un papel, se tratará de una mezcla sustractiva (del blanco del papel se sustraen las partes de la reflexión que no interesan o no se reflejan cuando es iluminado externamente).

La colorimetría basa su estructura teórica en tres principios fundamentales conocidos como las leyes de Grassmann, que hizo públicas en un artículo



DISTRIBUCION de la energía radiante espectral relativa al valor de 550 nanómetros en función de la longitud de onda para el espectro solar típico en el hemisferio norte con componente celeste incluida (*trazo lleno*), luz de tungsteno (*trazo de segmentos*) y un tubo fluorescente común (*trazo de puntos*).

editado en 1853. El primero de ellos es el ya enunciado, cuya formalización matemática es la siguiente:

$$c[C] \equiv r[R] + v[V] + a[A]$$

donde $[C]$ es la sensación de color, $[R]$, $[V]$ y $[A]$ corresponden a las sensaciones producidas por cada uno de los primarios r, v y a; c indica la magnitud física (por ejemplo, flujo radiante) de cada uno de los términos y el símbolo $\equiv$ indica equivalente.

El segundo principio expresa que, si se multiplica la equivalencia descrita por la ecuación anterior por un factor constante (K) en ambos términos, la equivalencia persiste. Esto significa en la práctica que la igualdad de color permanece constante independientemente de la magnitud física de los estímulos.

El tercer y último principio es el de aditividad, y se expresa diciendo que, si existen dos colores diferentes, son igualados por proporciones diferentes de los mismos primarios y la suma de ambos colores será equivalente a la suma de los componentes de cada uno de los primarios. Este tercer principio constituye la base fundamental de toda la colorimetría, ya que permite descomponer el espectro de la luz en fracciones simples de longitud de onda y sumar las contribuciones de cada una de ellas integrando el resultado final.

Si $P(\lambda)$ es la distribución espectral de determinada luz que llega al ojo y r_λ, v_λ, a_λ son las componentes tricromáticas por unidad de flujo radiante (conocida también como valores triestímulos reales del espectro equienergético) se puede escribir

$$R = \Sigma P(\lambda) \cdot r_\lambda \cdot \Delta\lambda$$

$$V = \Sigma P(\lambda) \cdot v_\lambda \cdot \Delta\lambda$$

$$A = \Sigma P(\lambda) \cdot a_\lambda \cdot \Delta\lambda$$

donde R, V y A son los valores triestímulos del color de la radiación en cuestión (o del objeto bajo una iluminación determinada). Por razones de conveniencia, la Comisión Internacional de Energía (CIE) decidió, en 1931, la adopción de tres primarios X, Y, Z –que son imaginarios– en reemplazo de los primarios reales R, V y A. Estos tres primarios imaginarios fueron elegidos de tal forma que uno y sólo uno de ellos es portador de toda la información concerniente a la luminosidad del color. La luminosidad de un color es correlativa al atributo psicológico de la claridad. Es decir: cuánta luz refleja una superficie dada e independiente de su cromaticidad. Se puede evaluar este factor comparando el color en cuestión con una escala de grises: cuanto más claros, tanta más luz reflejará. Así, por ejemplo, un rosa y un celeste pueden reflejar la misma cantidad de luz, su diferencia será tan sólo de cromaticidad.

La cromaticidad tiene dos componentes: el tono (la variación que caracteriza al color como rojo, naranja, amarillo, azul, verde, violeta, púrpura, etcétera) y la saturación, cuyo concepto es más difícil de entender. Si, por ejemplo, se realiza una mezcla de pintura, una roja y otra blanca, distintas proporciones darán lugar a una serie de colores que irán del rojo al blanco pasando por los rosados. Los colores a partir del rojo, manteniendo el tono, se irán desaturando y, además, aclarando; si también se agrega negro en una porción tal que la luminosidad del color resultante sea constante, el único factor de variación será la saturación, que será menor cuanto menor sea la proporción del rojo inicial en la mezcla total.

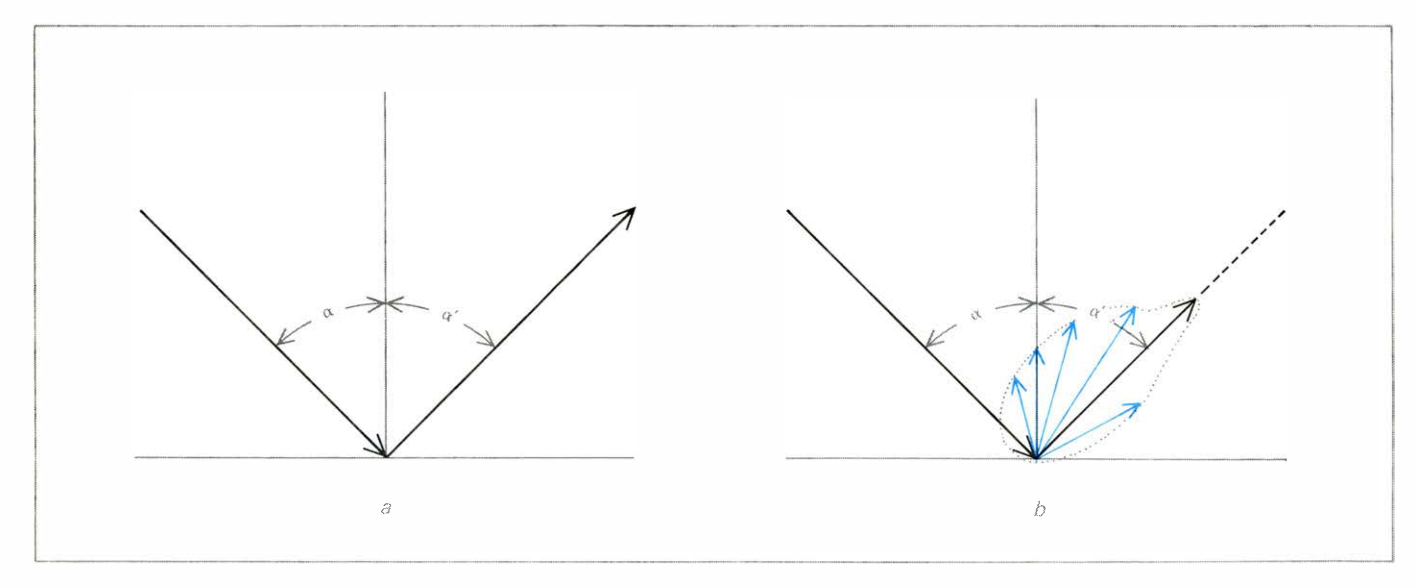
De lo dicho se deduce que entre dos colores similares puede haber diferencias de luminosidad, de tono y de saturación. Matemáticamente, al medir el color lo que se calcula son sus valores triestímulos X, Y, y Z. Por lo tanto, cualquier diferencia deberá manifestarse como un, ΔX, ΔY o $\Delta Z \neq 0$, donde ΔX, ΔY y ΔZ son las diferencias entre cada uno de los valores triestímulos de los colores en cuestión. El hecho de poder separar la luminosidad del color de su cromaticidad en el sistema elegido permite representarla en un espacio de sólo dos dimensiones. A esa representación se la conoce como espacio cromático CIE.

Para lograr esta representación se han representado las funciones $x_\lambda = \bar{x}_\lambda / (\bar{x}_\lambda + \bar{y}_\lambda + \bar{z}_\lambda)$ e $y_\lambda = y_\lambda / (\bar{x}_\lambda + \bar{y}_\lambda + \bar{z}_\lambda)$ donde $\bar{x}_\lambda$, $\bar{y}_\lambda$, $\bar{z}_\lambda$ son los valores triestímulos del espectro equienergético basado en los primarios imaginarios $[X]$, $[Y]$, $[Z]$ del CIE.

El trazo curvo de la ilustración de la página 12 representa los colores espectrales puros cuyas coordenadas han sido determinadas a través de sencillas ecuaciones. La recta que une los dos extremos cierra el espacio con las combinaciones del extremo violeta y el rojo (dando lugar a los púrpuras). Todos los colores tienen su representación cromática en esta área. Como dijimos anteriormente, al separar la cromaticidad de la luminosidad de un color, podemos representar el espectro cromático (o la cromaticidad) en el espacio CIE. Como consecuencia de ello el aspecto luminoso de un color se representa con un vector perpendicular al plano de cromaticidad cuyo módulo tiene el valor de Y. Por lo dicho, un color, en términos físicos, puede representarse tanto por los valores triestímulos X, Y, Z, como por las coordenadas cromáticas x, y y su luminosidad Y.

Ahora bien, una vez que se ha representado un color en este espacio mediante un punto en el mismo, otro color que sea similar a éste se representará con otro punto que estará más o menos próximo al anterior. Lo primero que a uno se le ocurriría sería medir la distancia entre esos dos puntos y tomar esa magnitud como expresión representativa de la diferencia entre ambos colores. Sin embargo, esto no es posible por razones puramente perceptivas: dos pares de colores que difieran entre sí lo mismo en el plano diagrama de cromaticidad del CIE no representan generalmente diferencias perceptuales iguales.

Una demostración de ello fue el estudio realizado por David L. MacAdam en 1942, quien midió el desvío medio normal de la dispersión de 25 diferentes colores en el espacio cromático. Esa dispersión dio lugar a elipses alrededor de los puntos representativos de cada color. Si se compara la ilustración de la página 13 con la de la 12 se puede ver que, en este plano, las diferencias de longitud no representan las mismas diferencias perceptuales; por ejemplo, la distancia entre dos colores que se ven iguales en el verde es mucho mayor que las distancias equivalentes en el rojo o en el azul. Y, lo que es más complicado aún, el valor consignado depende de la orientación; en otras palabras, perceptiblemente se detectan mejor diferencias,



ESQUEMA REPRESENTATIVO, a la izquierda, de la reflexión especular ideal en la que toda radiación incidente sobre la superficie (con un ángulo α respecto de la normal a ésta) es reflejada con un angulo α' (α' = α). A la derecha, se muestra la reflexión difusa más la especular de una superficie semimate en función del ángulo. Para un ángulo igual al de incidencia hay un aumento correspondiente a la componente especular.

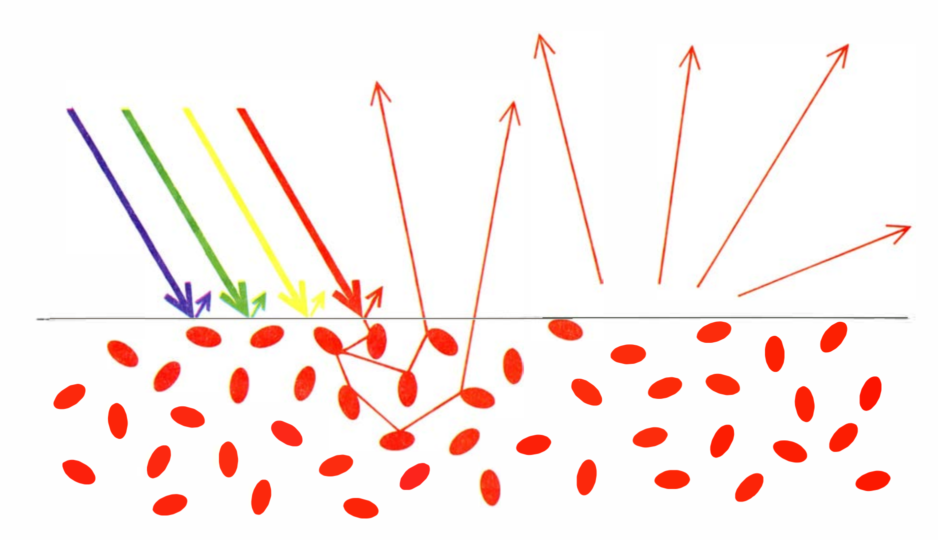
verbigracia, de tono que de saturación, o la inversa.

Este problema ha sido y es motivo de estudio de muchos expertos, y el CIE, a través del trabajo del Comité Técnico TC-1.3 de colorimetría, ha encauzado las investigaciones en la búsqueda de una geometría más adecuada. En términos matemáticos se dice que el espacio no es rectangular. A lo largo de los años se han propuesto distintas fórmulas y transformaciones que sería muy largo enumerar aquí. Algunas han sido lineales, otras no. Recientemente, el CIE ha propuesto dos alternativas como posibilidades reales. Ellas son las CIELUV y CIELAB. La primera corresponde a una transformación lineal del diagrama cromático del CIE. La segunda a una transformación no lineal. En ambos casos, las diferencias se evalúan según la clásica ecuación de distancia entre dos puntos donde se agrega la diferencia en luminosidad.

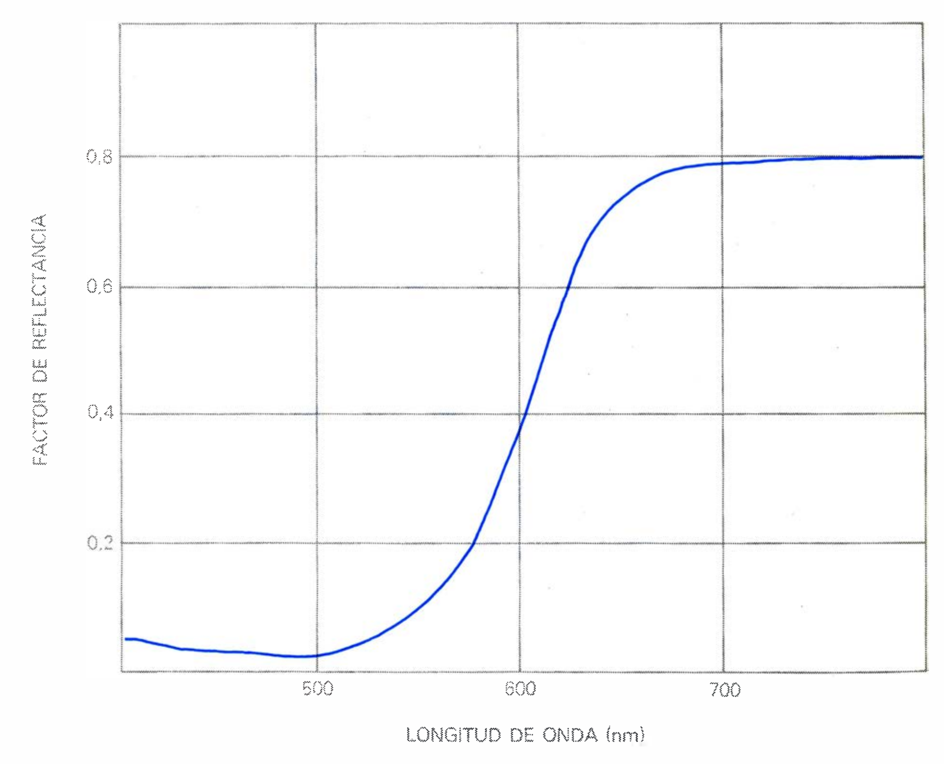
El problema de la correlación entre la percepción de los observadores y lo que miden las fórmulas descritas es tarea de investigación actual. Para ello se han planeado diversos experimentos que están dirigidos a verificar si las evaluaciones matemáticas determinadas según las transformaciones CIELUV y CIELAB predicen correctamente lo que los observadores ven. Existen diversos modos de realizar esta experiencia; uno de ellos es poseer colorímetros visuales donde el observador, por ejemplo, debe ajustar colores en distintas áreas contiguas que representan la misma diferencia perceptible en una dirección prefijada. Este tipo de instrumental existe en el Consejo Nacional de Investigación del Canadá. La ilustración superior de la página 14 muestra la disposición del campo visual.

El observador que tiene en el campo visual siete hexágonos, siendo fijo el color del centro, es instruido para que en los hexágonos circundantes modifique el color central en una dirección dada; así se modifica en pasos equivalentes y manteniendo en todos los casos la misma luminancia (o sea igual luminosidad) siguiendo variaciones en los hexágonos opuestos que corresponden a los mecanismos antagónicos de la visión de los distintos colores, siendo ellos: rojo $[R] = -$verde $[V]$, azul $[A] = -$amarillo $[Y]$ y saturado $[C] = -$desaturado $[D]$;

Luego, por algún método físico, cuando el observador ha logrado una solución satisfactoria, se miden las variaciones registradas y el ΔC de cada una de las observaciones comparándose entre sí y entre valores obtenidos con distintos colores. Otro método para verificar el comportamiento de estas fórmulas ha sido el empleo de técnicas psicofísicas (comparación por pares, estimación de la magnitud.) Los métodos consisten en pedir a los observadores que evalúen cualitativa o cuantitativamente muestras que se le presentan (en general de a pares o con una muestra testigo), pidiéndole un juicio de valor orientado a dar una respuesta sobre el problema particular que interese. Las muestras se preparan adecuadamente y conforman escalas de color donde se varía, en general, y en lo posible, sólo un parámetro (ya sea saturación, el tono o la luminosidad), estableciéndose de este modo métricas perceptuales. Estas se comparan estadísticamente con las predicciones de las fórmulas sometidas a pruebas.



EFECTO DE ABSORCION Y DIFUSION DE LA RADIACION electromagnética de un medio difusor (o túrbido). En él incide una radiación con componentes de distintas longitudes de onda representados en la figura por distintos colores (azul, verde, amarillo y rojo). Dado que las partículas en el medio (por ejemplo, pigmentos) son rojas, la radiación de este color es la única que se refleja dándole el aspecto cromático de rojo al objeto. Además, las direcciones reflejadas no tienen una única dirección debido a que hay interacción entre las partículas; se producen, pues, múltiples reflexiones y se origina como resultado una mayor o menor componente difusa (distintas direcciones de la especular).



CURVA CARACTERISTICA DEL FACTOR DE REFLEXION de un material rojo. En abcisas se representa la longitud de onda (dada en nanómetros) y, en ordenadas cartesianas, el valor relativo de la reflexión respecto de un difusor perfecto iluminado en igual forma que el material en cuestión.

Otros métodos empleados son los derivados del control estadístico de calidad, basados en criterios de pasa o no pasa. El problema es naturalmente complejo, debido a que las condiciones de evaluación no han sido constantes, ni tampoco el tamaño de las diferencias o las variaciones de las muestras. Por lo que los resultados aparecen, en ocasiones, como contradictorios. En nuestro laboratorio se ha estudiado el comportamiento de 13 fórmulas distintas en el caso de evaluación de diferencias de color medianas o grandes. Los resultados mostraron una aceptable correlación de la fórmula CIELAB con los valores experimentales; sin embargo, algunos de éstos indican un particular comportamiento del espacio y una dependencia significativa del modo de observación.

En este punto, hay que aclarar qué significa un modo de observación y qué importancia tiene. Para ello debemos retroceder hasta el momento en que la radiación llega al ojo y es detectada por el sistema visual. Hasta aquí hemos dejado implícita la idea de que a un tipo de estímulo corresponde una y sólo una respuesta sensorial. Lo que no se dijo, y debe ser aclarado ahora, es que puede haber muchos estímulos diferentes que produzcan la misma respuesta, y puede también ocurrir que dos estímulos de color idénticos produzcan distintas sensaciones dependiendo ello de la forma en que se siente el estímulo.

Sin querer ser exhaustivos, los factores que determinan las condiciones de contorno de una percepción de color son: nivel de iluminación, campo visual, contraste, separación entre muestras, fondo, tipo de luz empleada, tiempo y otros aspectos (como textura, brillo, forma, etcétera). Lo natural es tratar de fijar estos parámetros para realizar experiencias cuyos resultados sean comparables. Ello implica, sin embargo, restricciones que exceden muchas veces la práctica cotidiana. Como el último objetivo es encontrar un medio idóneo para ser empleado en ésta, las tolerancias en la aplicabilidad de una fórmula dada deben ser lo más amplias posible; y se requiere estudiar cómo varía el comportamiento cuando modificamos algunos de los parámetros enumerados anteriormente. Se hace necesario, pues, describir, aunque sea de un modo sumario, en qué consisten estos parámetros de luminancias.

Es conocido que el ojo ve desde niveles muy bajos hasta muy elevados. El ámbito es de aproximadamente de 10^{-6} a 10^{6} candelas por metro cuadrado, o sea aproximadamente 12 unidades logarítmicas. (La candela es la unidad fundamental de intensidad lumínica. Equivale a la intensidad lumínica de 1/600.000 metros cuadrados de la superficie de un cuerpo negro a una temperatura de 2040 grados kelvin y a una presión de 101.325 pascal.) Esta versatilidad se logra mediante la activación de dos tipos de detectores; los bastones (los más sensibles) y los conos. Estos últimos son los principales responsables de la visión del color, por lo que el nivel debe ser suficientemente elevado como para que el observador lo aprecie claramente. A este nivel se le ha llamado fotópico (por encima de una candela dividido por metro cuadrado).

El segundo aspecto a considerar es el campo visual, es decir, el ángulo sólido determinado por el objeto de color. Si la estructura fisiológica y nerviosa de la retina del ojo fuera uniforme, esto no entrañaría ningún problema, pues probablemente la respuesta sensorial sería independiente de la zona de la retina afectada. Pero esta estructura difiere notablemente según sea el área considerada. Si entendemos el ojo como un sistema óptico, en el centro óptico de la retina existen cambios morfológicos considerables, adaptándolo particularmente para la visión fotópica y una máxima agudeza visual. En un área reducida que no va más allá de los cuatro grados alrededor del eje hipotético que pasara por el centro de la fóvea, sólo existen conos. Por tanto, esta zona no sólo es responsable de la mayor discriminación (o agudeza visual), sino también de la mayor información relativa al color.

Se ha probado que, a medida que uno se desplaza a zonas contiguas de la fóvea, la percepción del color se reduce volviéndose el observador dicrómata (pierde una de las sensibilidades cromáticas) primero y, en algunos casos, acrómata cuando la zona afectada es la periferia. Por otra parte, en la práctica muchas veces es necesario observar campos visuales más amplios (hasta 10 grados), por lo que el CIE que en 1931 había recomendado las funciones r_λ, v_λ, a_λ (o sus equivalentes x_λ, y_λ, z_λ) para un campo visual de dos grados, en 1963 re-

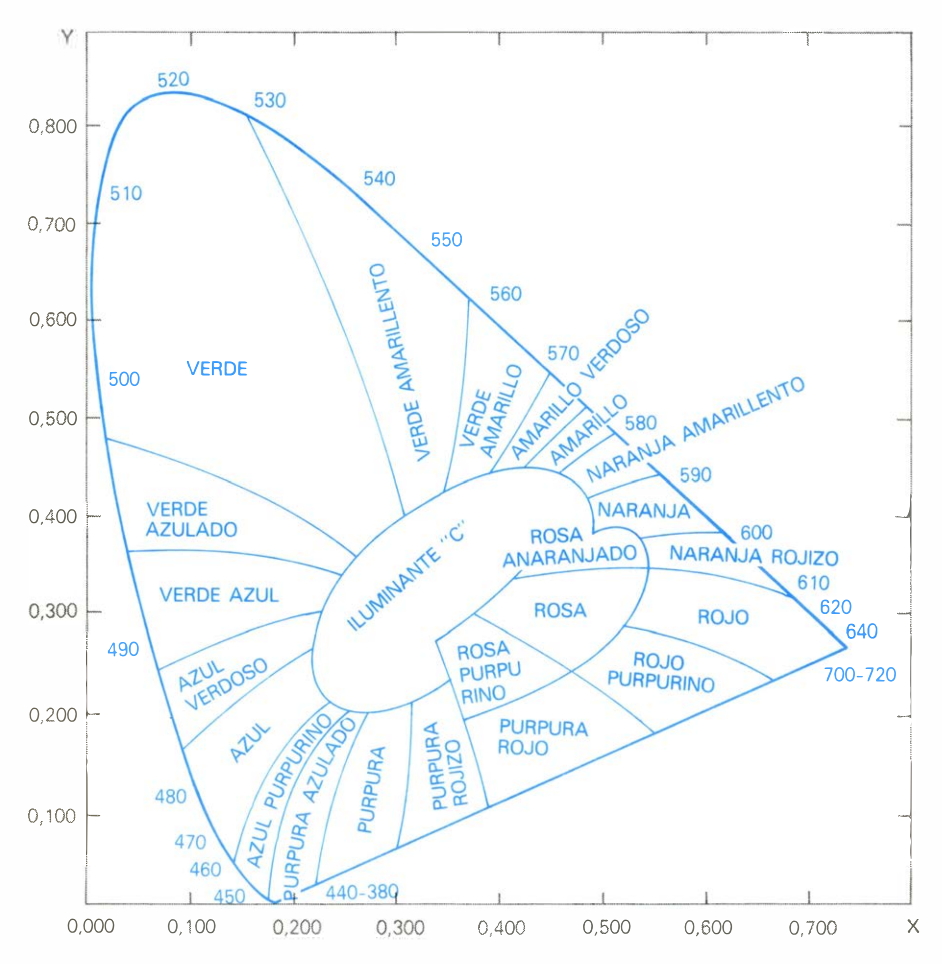


DIAGRAMA CROMATICO CIE donde se dibujan las áreas aproximadas correspondientes a cada zona de un determinado color. La línea curva representa los colores de las radiaciones espectrales puras. La recta inferior que cierra esta curva se llama línea de púrpuras, y representa todas las posibles combinaciones de las radiaciones extremas del ámbito visible. En el área encerrada en este espacio tienen su representación todos los colores reales. Este es un plano de cromaticidad, por lo que no se representa en él la luminosidad del color; ésta puede indicarse por un vector perpendicular al plano, cuyo módulo tenga el valor correspondiente a *Y* y esté aplicado al punto *x, y* representativo del color.

comendó nuevas funciones que son válidas cuando el campo visual es mayor de cuatro grados y menor de 10 grados. Uno de los problemas que se presentan en las mediciones de estas funciones fue la participación (o intrusión como se la dio en llamar) en la misma de los bastones. Otro de los problemas fue la percepción de lo que se denomina "mancha de Maxwell" y que se debe a la *mácula lutea*, que es una capa pigmentada que se puede detectar alrededor de la fóvea y que está extendida hasta los 5 grados, aproximadamente. Como parte de estos fenómenos se observó una no-aditividad de las nuevas funciones, no-aditividad que significa el incumplimiento de la tercera ley de Grassmann que mencionáramos anteriormente. Por tanto, se hizo necesario especificar claramente cuál era el campo visual en las experiencias de evaluación de diferencias de color a fin de evitar problemas que surgirían del empleo de diferentes condiciones. Dificultades, por otra parte, casi insalvables, debido a la necesidad de que las transformaciones del espacio cromático fueran válidas para ambas posibilidades (foveal o parafoveal).

El tercer factor a tener en cuenta es el contraste. Es conocido el hecho de que una muestra o un mismo objeto de color no parece tener el mismo color si lo observamos contra fondos cromáticos diferentes, induciendo modificaciones que dependen de la cromaticidad del fondo. Además, la habilidad de los observadores para detectar diferencias está en relación inversa con la separación de las muestras y del contraste luminoso respecto del fondo, por lo que siempre se trata de que las muestras se vean contra un fondo neutro de claridad (luminancia), similar al de las muestras a evaluar y que las mismas estén lo más juntas posible.

El tipo de luz empleado en la igualación es de vital importancia. No es lo mismo la luz natural, una lámpara incandescente o un tubo fluorescente, ya que todos ellos tienen distribuciones espectrales características $E(\lambda)$, que difieren entre sí y, por tanto, producirán modificaciones en la percepción de los colores. Existen casos donde dos muestras tienen el mismo color bajo un tipo determinado de luz o iluminante y son diferentes bajo otro iluminante. (Este fenómeno recibe el nombre de metamerismo.) Por lo que el CIE recomienda distintos tipos de luz caracterizados como "*C*" o "*D*65", que corresponden a la luz solar media; la luz *C* no tiene especificada la componente ultravioleta mientras que la luz *D*65, sí.

El siguiente factor a tener en cuenta es el tiempo en que el observador realiza la evaluación, ya que depende de éste la precisión de los mismos. El CIE no ha recomendado un valor fijo, pero pide que sea suficiente para que el observador tome la decisión sin vacilar. El último factor es el modo de apariencia, pues el color es sólo uno de estos aspectos. A fin de que las evaluaciones no estén influenciadas por otros efectos diferentes del color, se ha recomendado que las muestras no posean textura ni que pueda verse su brillo, así como que sean superficies de la misma forma, sin que lleguen a diferenciarse entre sí por estos factores. En el trabajo realizado en nuestro laboratorio basado en experiencias llevadas a cabo por la arquitecto Fago de Mattiello en el Laboratorio de Investigaciones Sensoriales, se estudiaron los efectos de la variación de campo visual y fondo observado (si iluminado o no) en la geometría del espacio cromático. De algún modo, los resultados mostraron una significativa mejora estadística para un fondo iluminado respecto de otro no iluminado. Además, esta diferencia se acentuó para campos visuales reducidos (de 42 minutos) respecto del de cuatro grados empleado en la mayoría de las evaluaciones.

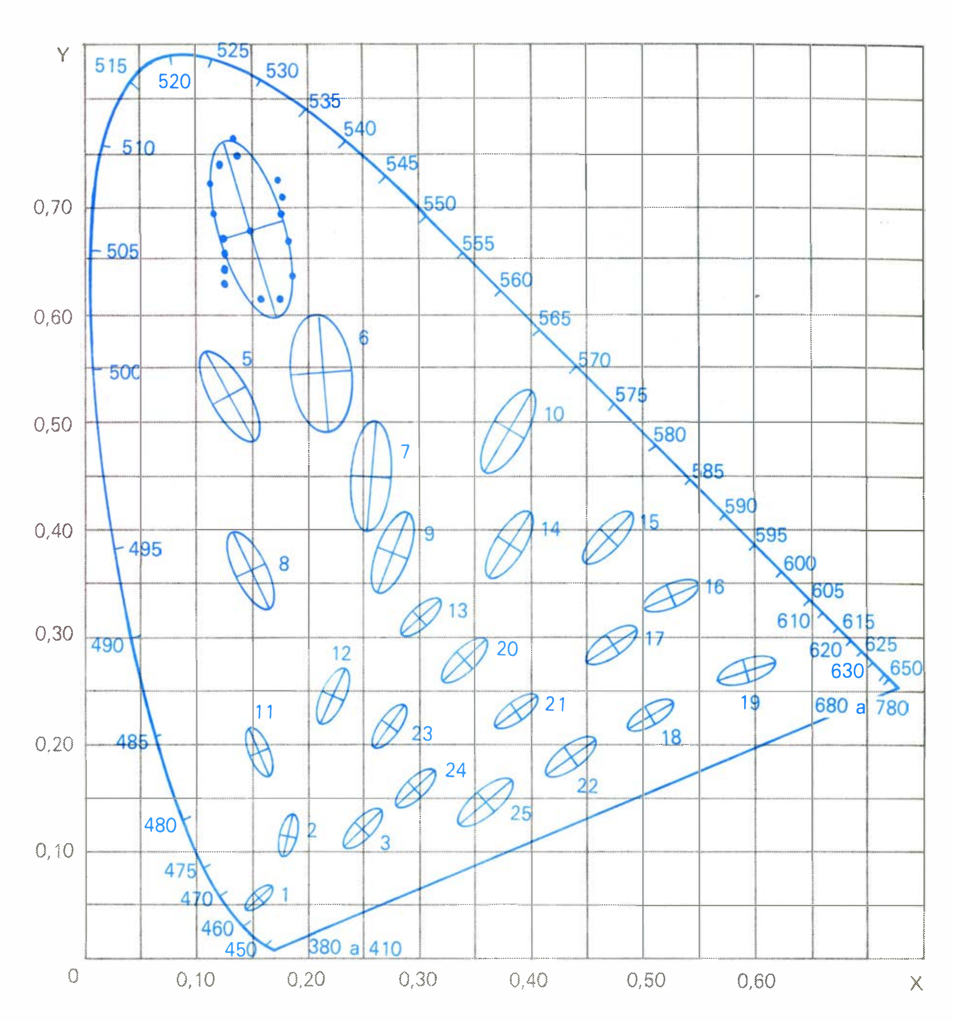
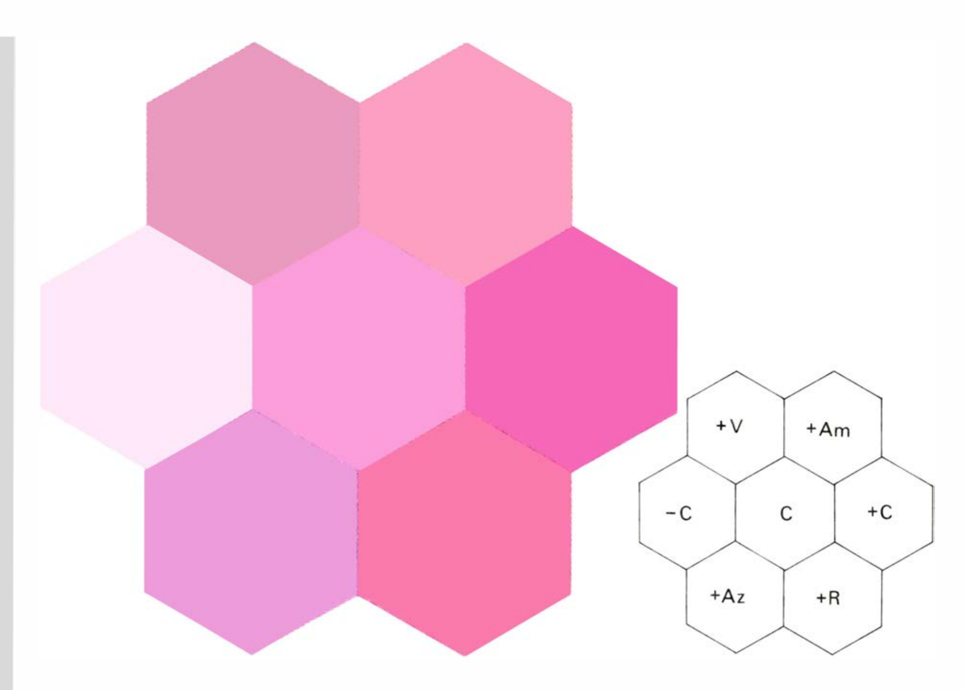


DIAGRAMA CROMATICO CIE donde se dibujan las elipses correspondiente a las mínimas diferencias perceptibles para 25 colores según el experimento de MacAdam. Comparada con la ilustración anterior, aquí las diferencias de longitud no representan las mismas diferencias perceptuales.

Este comportamiento no es uniforme en todas las fórmulas y en particular la CIELAB sigue un proceso bastante diferenciado del CIELUV. En lo esencial, el CIELUV es una transformación lineal del espacio cromático CIE, esto es, líneas rectas en el diagrama *x*, y del CIE, se representan con líneas rectas en el u^*, v^*. No es así en cambio, en el a^*, b^* (otro sistema, propuesto por Glasser) donde suelen representarse con líneas curvas.

¿Cuál es la aplicabilidad en la práctica, de una u otra fórmula? No está definido aún; pareciera que ello sólo depende del particular interés de cada usuario. La realidad muestra que, en principio, la fórmula CIELAB y el espacio cromático por ella definido es más uniforme que el CIELUV. Es reconocido que ninguna transformación lineal del espacio cromático del CIE puede solucionar sus asimetrías y que es necesaria una transformación de orden superior para mejorar la uniformidad del

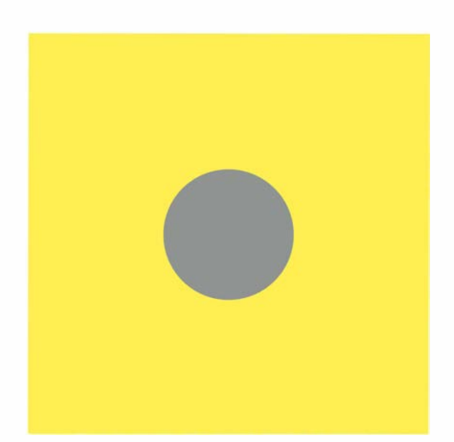


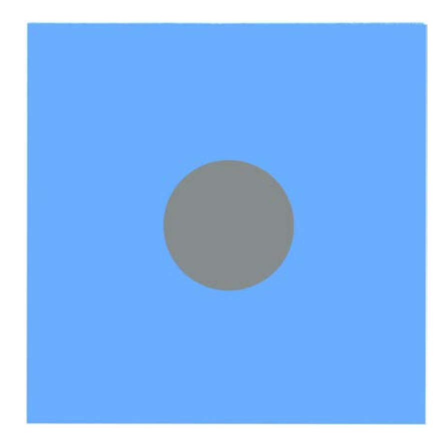
CAMPO VISUAL del colorímetro del NRC del Canadá. En él se representan las variaciones cromáticas respecto de un color de prueba (ubicado en el hexágono central). En pasos perceptibles iguales, las variaciones se realizan agregando amarillo (+ Am) sumando rojo (+ R), restando amarillo (−Am) que es equivalente a sumar azul (+ Az), disminuyendo el rojo (−R), que es equivalente a agregar verde (+ V), más saturado y menos saturado (o más saturado). (La reproducción de los colores es sólo aproximada, pues en el colorímetro se suman luces; pero en esta producción se emplean principios sustractivos).

mismo; quizá sea por ello por lo que los resultados experimentales muestran una mayor correlación para el CIELAB con las evaluaciones subjetivas. Por otra parte, dado el tamaño de las diferencias evaluadas (medianas y grandes en nuestro experimento) y la concordancia con otros experimentos donde el tamaño de las mismas era mucho menor, cabe esperar en esta fórmula en la mayoría de los casos prácticos. Si esto ocurre, la misma práctica hará caer en desuso la fórmula CIELUV.

No deseo terminar el trabajo sin destacar algo esencial para esta tarea de investigación: las dificultades de encontrar explicaciones idóneas (esto es: matemáticamente correctas) que describan el comportamiento humano, que en este caso particular se refiere a su sentido de la visión y particularizando aún más a la visión de los colores.

La definición de una estructura físico-matemática que describa el funcionamiento del mecanismo visual es todavía una utopía. La enorme complejidad y la ignorancia de muchos de los procesos implicados −pese al enorme caudal de información que se posee en la actualidad− hace que las soluciones sean tan sólo aproximaciones, útiles en la medida que sirven para resolver algunos problemas prácticos (a veces muy importantes y de gran trascendencia económica), pero que no constituyen otra cosa que aproximaciones a una verdad que, hoy por hoy, es todavía inasible.





EFECTOS DE CONTRASTE CROMATICO. Las áreas centrales son de grises iguales, pero que aparentan ser diferentes por el contorno cromático que los rodea. El contraste condiciona la percepción.

Conductores de cadena lineal

Ciertos materiales con estructura lineal o en forma de columna conducen bien la electricidad únicamente a lo largo de un eje determinado. La organización monodimensional de estos materiales configura sus propiedades electrónicas

Arthur J. Epstein y Joel S. Miller

La experiencia con las sustancias ordinarias nos enseña a asociar ciertas propiedades con las diferentes clases de materiales. Por ejemplo, metales como el cobre, el hierro y el níquel, al igual que las aleaciones metálicas, son buenos conductores de la electricidad. Los materiales iónicos tales como la sal común, los sólidos moleculares como el azúcar y los polímeros sólidos como el teflón y el poliestireno suelen ser malos conductores en general. Sin embargo, recientemente se han obtenido ciertos materiales que no se comportan en la forma que cabría esperar a priori. Se trata de cristales iónicos, sólidos moleculares y polímeros cuya estructura electrónica presenta cierta similitud con la de los metales y, en consecuencia, algunas de sus propiedades son también análogas a las de éstos. En particular, dichos materiales son buenos conductores de la electricidad.

La propiedad más importante de estos materiales es su anisotropía: su conductividad es diferente cuando se mide a lo largo de direcciones distintas en el sólido. En la mayoría de los casos existe un eje a lo largo del cual la conductividad es elevada y, en algunas ocasiones, dicha conductividad resulta incluso comparable a la de los elementos metálicos. En una dirección perpendicular a dicho eje, la conductividad puede llegar a ser menor que la anterior en un factor del orden de 100.000. Esta anisotropía es un reflejo de la estructura interna de dichos materiales, constituidos por muchas cadenas o pilas de moléculas. La conducción de electricidad tan sólo tiene lugar a lo largo de estas cadenas o pilas. Las sustancias a las que nos estamos refiriendo se conocen con el nombre de sólidos cuasi-monodimensionales o materiales de cadena lineal.

Aunque todas las propiedades eléctricas, magnéticas, ópticas, estructurales y químicas de estos materiales son interesantes, nosotros nos ocuparemos fundamentalmente de su conductividad eléctrica. En algunos casos, el mecanismo de la conducción en estos materiales es muy similar al de los elementos metálicos, o por lo menos dicho mecanismo es análogo al de los metales por encima de una determinada temperatura umbral. Para temperaturas inferiores a ésta, los electrones de estos materiales sufren una transición de fase y los materiales se convierten en semiconductores. Otros sólidos de cadena lineal son semiconductores o aislantes dentro de intervalos de temperatura muy amplios. Finalmente, existen otros materiales cuyas propiedades sólo pueden describirse de forma adecuada mediante modelos de conducción más complejos.

No todos los sólidos de cadena lineal de alta conductividad han sido descubiertos en fechas recientes. En el año 1842, W. Knop preparó una sal de tetracianuro de platino, un grupo químico en el que cuatro iones cianuro (CN^-) rodean un ion de platino. Esta sal es soluble en agua y su color y brillo son análogos a los del oro o bronce. En 1910 Frank Playfair Burt, del University College de Londres, sintetizó un polímero inorgánico del nitruro de azufre que consistía en átomos de azufre y nitrógeno enlazados en forma de cadenas largas. La fórmula química de este compuesto es $(SN)_x$ donde x representa un número grande e indeterminado. Estas dos sustancias son sólidos de cadena lineal cuya conductividad es elevada. En el $(SN)_x$ las largas filas polímeras se alinean paralelamente entre sí. El grupo del tetracianuro de platino tiene la forma de un disco plano, pero en la sal preparada por Knop los discos se apilan formando grupos paralelos. Esta estructura fue identificada por vez primera en el año 1964 por Klaus Krogmann, de la Universidad de Stuttgart. Otras sales de tetracianuro de platino poseen una estructura similar a ésta; nosotros nos referiremos a todas ellas utilizando la abreviatura genérica TCP.

Aunque el TCP y el $(SN)_x$ se conocen desde hace muchos años, sus propiedades no se estudiaron con detenimiento hasta que se inició la década de los años 1970. Por aquel entonces, se habían preparado ya otros sistemas de cadena lineal. En 1960 un grupo de investigadores de la E. I. Du Pont de Nemours and Company sintetizó e identificó una molécula particularmente importante. Su nombre formal es 7,7,8,8 – tetracianuro – p – quinodimetano y suele escribirse abreviadamente como TCNQ. Se trata también de una molécula plana, pero, al contrario del TCP, el TCNQ es una sustancia orgánica constituida exclusivamente por átomos de carbono, nitrógeno e hidrógeno. Muchas de las sales formadas por la combinación del TCNQ con otros átomos o moléculas son sólidos de cadena lineal.

A partir del año 1964, el estudio de los sólidos de cadena lineal recibió un importante estímulo gracias a los trabajos de W.A. Little, de la Universidad de Stanford. Este investigador sugirió que si pudiéramos producir un material de cadena lineal en forma tal que cumpliera unas determinadas especificaciones, dicho material presentaría el fenómeno de la superconductividad no sólo para temperaturas bajas, como sucede con los superconductores propiamente dichos, sino también para temperatura ambiente. La estructura que Little proponía estaba formada por un espinazo conductor alrededor del cual se sitúan otros grupos que estabilizan el transporte superconductor de los electrones. La existencia de un superconductor a temperatura ambiente tendría una gran importancia tecnológica y la idea de Little alentó los esfuerzos realizados para sintetizar tales estructuras. Sin embargo, todavía no se ha descubierto ningún superconductor a temperatura ambiente y la propuesta de Little sigue siendo motivo de controversia.

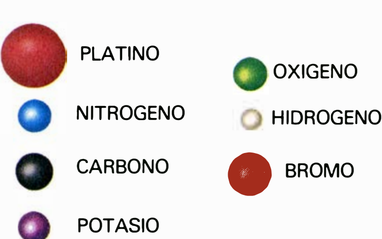
En el año 1973, Alan J. Heeger, Anthony F. Garito y sus colaboradores, de

la Universidad de Pennsylvania, observaron una conductividad muy alta en una sal formada por el TCNQ y otra molécula orgánica, el tetratiofulvaleno, a la que designaremos abreviadamente como TTF. Para temperaturas relativamente bajas (entre 50 y 60 grados Kelvin), estos investigadores encontraron que la conductividad del (TTF) (TCNQ) es del orden de la conductividad del cobre a temperatura ambiente. En un principio se pensó que este estado altamente conductor podría significar el acceso a la superconductividad, pero las cosas resultaron de otro modo.

Sin embargo, lo que se encontró en realidad no resultaba menos interesante. La estructura en columna de los conductores anisótropos es esencial para su naturaleza e influye decisivamente en sus propiedades. Los mecanismos de conducción observados en estas sustancias difieren del mecanismo que opera en un elemento metálico de tres dimensiones. Ciertas transiciones electrónicas de fase, que convierten a un metal en un semiconductor, son características de los materiales monodimensionales. El interés de estos fenómenos radica en lo que son capaces de revelar acerca de la física y la química del estado sólido, además de las aplicaciones tecnológicas de los propios materiales.

La conductividad eléctrica es una propiedad, de enorme importancia, de la materia. La unidad de medida de esta magnitud física es el mho dividido por centímetro, que es la inversa de la unidad de resistividad, el ohm multiplicado por centímetro (la palabra "mho" fue introducida por Lord Kelvin, en contraposición al "ohm", la unidad que debe su origen a Georg Simon Ohm). Las conductividades de los mejores aislantes, tales como el teflón y el poliestireno, son del orden de 10^{-18} mho/centímetro; las de los mejores conductores a temperatura ambiente, el cobre y la plata, son del orden de 10^{6} mho/centímetro. La distinta conductividad de algunos materiales de cadena lineal (medida a temperatura ambiente en la dirección del eje favorable) se sitúa entre un centenar y algunos miles de mho/centímetro.

La conductividad de un sólido viene determinada por su estructura electrónica. La distribución de los electrones en la proximidad de un átomo se describe mediante un sistema de orbitales, cada uno de los cuales posee una forma y un tamaño que le son característicos. En un sólido de alta conductividad los orbitales de los átomos o moléculas adyacentes se solapan entre sí, de forma tal que los electrones pueden moverse con facilidad de un lugar a otro a lo largo de la red. En los conductores de cadena lineal el solapamiento de los orbitales es importante tan sólo a lo largo de un eje dado. Por ejemplo, en los sistemas constituidos por tetracianuro de platino, los orbitales en cuestión son orbitales particulares del ion platino, conocidos como orbitales d_{z^2}; estos orbitales se extienden a ambos



LAS MOLECULAS APILADAS UNA SOBRE OTRA proporcionan el eje fundamental de la conducción en una sal de tetracianuro de platino (TCP). Cada unidad de TCP es un disco plano con un ion de platino en el centro. Los discos están apilados de forma tal que los orbitales (o nubes de electrones) de los iones de platino adyacentes de una cadena determinada se solapan entre sí. La conductividad es alta en la dirección a lo largo de dichas pilas, pero no en las direcciones perpendiculares a ésta. Además del tetraciamero de platino (TCP), el cristal al que nos estamos refiriendo incluye iones de bromo y potasio junto con moléculas de agua. Los diferentes tipos de átomos se identifican mediante colores diferentes, de acuerdo con el código que aparece a la izquierda de la figura. Las líneas discontinuas representan enlaces de hidrógeno. La fórmula de la sal es $K_2Pt(CN)_4Br_3 \cdot 3H_2O$. En la portada de este número de *Investigación y Ciencia* se reproduce este mismo cristal en una forma diferente (*véase página dos*).

lados de los planos del TCP. En una sal de TCP la separación entre dos iones de platino es de 2,88 angstrom (un angstrom equivale a 10^{-8} centímetros) y ello permite que solapen los orbitales d_z2 de los iones de platino que están situados en planos adyacentes. A título de comparación, diremos que la distancia interatómica en el platino metálico (un conductor excelente) es de 2,79 angstrom.

Como resultado del solapamiento de los citados orbitales, los electrones se deslocalizan a lo largo de las columnas que constituyen el TCP. El movimiento de electrones desde una columna a otra es mucho menos eficiente que el que corresponde a la dirección anterior. Una sal de TCP típica, que incorpora iones de potasio y bromo así como agua, tiene la fórmula $K_2Pt(CN)_4Br_{0,3} \cdot 3H_22O$. Hansrudi R. Zeller y sus colaboradores, del Centro de Investigación Brown-Boveri, han puesto de manifiesto que la conductividad de este material en una dirección paralela a las columnas de TCP es de unos 300 mho/centímetro. La conductividad en una dirección perpendicular a la anterior es 100.000 veces menor.

El (TTF) (TCNQ) está constituido por dos tipos de moléculas planas que forman columnas conductoras paralelas entre sí. Las moléculas de cada fila de columnas están inclinadas respecto al eje de dicha fila y las columnas adyacentes se disponen en forma parecida a la espina de un arenque. Dentro de las columnas de TTF y TCNQ los orbitales de cada molécula se solapan con los de las moléculas situadas encima y debajo de ella en la misma fila. Los orbitales en cuestión están asociados con la molécula como un todo y no con un átomo en particular. Existe poco contacto entre los orbitales de las moléculas situadas en filas diferentes y, por consiguiente, la eficiencia del transporte de cargas en cualquier dirección perpendicular a las filas es muy baja. La conductividad máxima del (TTF) (TCNQ) a temperatura ambiente, en una dirección paralela a las columnas, es mayor que la del TCP: del orden de 650 mho/centímetro. Sin embargo, la anisotropía es tan sólo unas 500 veces mayor que la que corresponde a la dirección menos favorable.

Muchos polímeros están constituidos por moléculas enlazadas químicamente para formar cadenas largas y, en consecuencia, presentan cierta anisotropía. No obstante, la mayoría de los polímeros son malos conductores de la electricidad. El $(SN)_x$ es una excepción: su conductividad a temperatura ambiente es de unos 2000 mho/centímetro. Esta elevada conductividad se debe al solapamiento de orbitales paralelos a las filas del polímero, facilitando así la conducción a lo largo de dichas filas. La conductividad en las direcciones perpendiculares a las filas en cuestión es relativamente alta (en otras palabras, la anisotropía es baja), poniendo de manifiesto con ello el hecho de que el solapamiento de los orbitales entre las diferentes cadenas es también importante. El $(SN)_x$ se comporta como un superconductor convencional de tres dimensiones para temperaturas inferiores a 0,3 grados Kelvin.

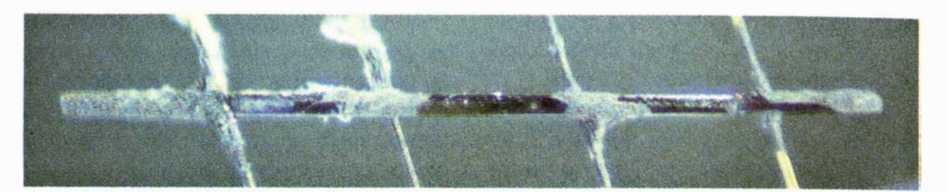
El poliacetileno, cuya fórmula es $(CH)_x$, es un polímero orgánico cuya conductividad puede ser alta. Este compuesto está constituido por una cadena continua de átomos de carbono, que se representa generalmente en forma tal que los enlaces dobles y simples que la integran se presentan alternados y en donde cada átomo de hidrógeno está enlazado a un átomo de carbono diferente. La adición de ciertas impurezas en concentraciones variables permite variar la conductividad de este compuesto dentro de un margen muy amplio, desde 10^{-9} mho partido por centímetro hasta más de 2000 mho/centímetro.

La sensibilidad de las estructuras lineales a los defectos y al desorden es muy grande. En un sólido tridimensional, un defecto puntual puede reducir la conductividad como resultado de la difusión de electrones que produce dicha imperfección, pero es incapaz de bloquear totalmente la conducción debido a que los electrones pueden dar un rodeo al llegar a él. Por el contrario, en un material de cadena lineal, una interrupción de la cadena puede detener totalmente la conducción a lo largo de la correspondiente fila. Evidentemente, si se rompe un número elevado de filas en algún punto a lo largo de toda su extensión, la conducción global se verá seriamente afectada.

Dado que los sistemas lineales son muy sensibles a la disrupción, su conductividad intrínseca se puede medir con precisión únicamente en el caso de que se trate de cristales puros y casi perfectos desde el punto de vista cristalográfico. Con el fin de preparar tales cristales se han puesto a punto ciertas

CRISTALES EN FORMA DE AGUJA de un material de cadena lineal cuya conductividad es elevada. Estos cristales fueron obtenidos mediante un proceso electroquímico en el Centro de Investigación Webster de la Xerox Corporation. El material en cuestión es una sal de potasio del TCP cuya fórmula es $K_{1,75}Pt(CN)_4 \cdot 1,5\ H_2O$. Su brillo metálico es característico de muchos conductores monodimensionales. La longitud de estos cristales es inferior a un milímetro. Las pilas de discos de TCP se disponen paralelas al eje de las agujas y, por tanto, la conductividad a lo largo de esta dirección es máxima.

técnicas especiales. Uno de estos métodos, desarrollados en el Centro de Investigación Webster de la Xerox Corporation, utiliza la propiedad a medir –la conductividad– para hacer crecer el cristal. Cuando una corriente eléctrica atraviesa una disolución de iones potasio e iones de tetracianuro de platino, sobre el ánodo se forman cristales de TCP. La conductividad de las muestras de la sal $K_{1,75}Pt(CN)_4 \cdot 1,5\ H_2O$, preparadas de este modo, y medida aquella medida a temperatura ambiente, es unas 100.000 veces mayor que la de los cristales cuyo crecimiento se ha logrado utilizando métodos más convencionales.



PARA MEDIR LA CONDUCTIVIDAD de un conductor de cadena lineal es necesario conectar cuatro terminales en un cristal de 1,5 milímetros de longitud. Los dos terminales más externos se utilizan para hacer pasar una corriente conocida a través del cristal; a continuación, se mide el voltaje originado en los terminales más internos. Dicho voltaje es proporcional a la resistencia eléctrica de la muestra. El inverso de esta resistencia es la conductancia. El valor de la conductividad depende también de las dimensiones de la muestra. El cristal que aparece representado en esta ilustración es una muestra de (NMP) (TCNQ); dicho cristal está montado sobre hilos de oro de 0,025 milímetros de diámetro.

Los monocristales de sólidos de cadena lineal tienen frecuentemente forma de aguja; las filas conductoras se sitúan en direcciones paralelas al eje de la aguja. En general, los cristales son bastante pequeños: suelen tener menos de un milímetro de longitud y unas centésimas de milímetro de grosor. Observados a la luz ordinaria, muchos conductores de cadena lineal presentan un brillo metálico. Las sustancias inorgánicas que incorporan al elemento platino entre sus constituyentes suelen tener el color del oro o del bronce, y lo mismo sucede con el $(SN)_x$. Las láminas de poliacetileno tienen el aspecto de hojas de aluminio, pero su textura se asemeja más bien a la de los plásticos utilizados para envolver alimentos. Los materiales orgánicos tales como el (TTF) (TCNQ) son generalmente negros.

Con el fin de explicar el mecanismo de la conducción eléctrica es necesario referirnos al estado en que se encuentran los electrones de un sólido. En un átomo aislado cada electrón ocupa un orbital determinado que posee cierta energía asociada con él. El orbital define la distribución de la carga del electrón en el espacio. Cada orbital puede estar ocupado a lo sumo por dos electrones. En un átomo en su estado fundamental, los dos primeros electrones se sitúan en el orbital de energía más baja y todos los demás electrones ocupan otros orbitales de energía superior a la de éste. Dado que todos sus electrones están confinados en orbitales, los niveles energéticos de un átomo son discretos; todos los estados de energías intermedias están prohibidos. Si un electrón cambia su energía, debe producirse un salto desde un nivel permitido a otro nivel también permitido.

Cuando dos átomos idénticos se sitúan lo suficientemente próximos entre sí como para que sus orbitales se solapen, cada nivel de energía se desdobla para crear dos nuevos niveles, uno por debajo del nivel original y el otro por encima de él. La magnitud del desdoblamiento viene determinada por la cuantía del solapamiento de los orbitales o, en otras palabras, por la separación entre los dos átomos considerados. Es importante señalar que el número total de electrones que puede poseer un sistema no se altera cuando los dos átomos se unen para formar una molécula. Mientras que antes de dicha unión había dos orbitales atómicos independientes de la misma energía, cada uno de los cuales podía alojar a dos electrones, después de ella tendremos dos orbitales moleculares con dos niveles energéticos diferentes y en cada uno de los cuales existirán dos electrones. Cuando se unen tres átomos se crean tres nuevos niveles de energía, uno por debajo del nivel original, otro por encima de él y el tercero situado entre los dos anteriores.

En un sólido cristalino tridimensional existen del orden de 10^{23} átomos con sus orbitales solapados. Cada nivel energético de los átomos originales se desdobla en un cierto número de niveles poco separados entre sí. De hecho, la separación entre los diferentes niveles energéticos adyacentes es tan pequeña que el conjunto de niveles se puede considerar como una banda continua en la que un electrón puede tomar cualquier energía que esté comprendida entre los extremos superior e inferior de la banda. Puesto que cada uno de los niveles que constituyen la banda solamente puede alojar un máximo de dos electrones, en un sólido constituido por n átomos cada banda estará ocupada a lo sumo por $2n$ electrones. Si se añaden más electrones, dichos electrones deberán alojarse en la banda desocupada de energía más baja y en algunas ocasiones esto sólo es posible mediante la aportación de una cantidad significativa de energía.

El grado de ocupación de las bandas de energía y la magnitud de los intervalos existentes entre ellas (bandas prohibidas) determinan la mayoría de las propiedades eléctricas de los sólidos. Se suele decir que los electrones que ocupan una banda están deslocalizados: estos electrones se pueden desplazar libremente por el interior del sólido. Sin embargo, para mantener una corriente eléctrica debe haber un movimiento neto de electrones en una determinada dirección. Si todos los niveles energéticos de una banda se llenan de forma continua desde el nivel más bajo disponible hasta el nivel más alto permitido no puede haber un movimiento neto de los electrones.

Las propiedades peculiares de los metales se deben precisamente a la presencia de una banda de energía parcialmente llena. Los electrones situados en los niveles de energía más alta de un metal pueden pasar fácilmente a ocupar los niveles vacantes, cuya energía supera a la de los anteriores en una cuantía infinitesimal. De este modo, siempre existen electrones disponibles para la conducción. Cuando los orbitales que se combinan para formar una banda no están totalmente ocupados, la banda en cuestión estará parcialmente ocupada. Por ejemplo, si todos los orbitales de un sólido están ocupados solamente por un electrón (en lugar del máximo de dos electrones), la correspondiente banda de energía en dicho sólido tendrá n electrones en lugar de $2n$. Los restantes niveles de la banda están disponibles para la conducción.

La conductividad de un metal resulta limitada por la difusión o dispersión de los electrones en movimiento; este fenómeno altera la dirección de dicho movimiento y tiende a restablecer un estado en el que el momento neto de todos los electrones es igual a cero. Para altas temperaturas, la causa principal de la difusión es la interacción de los electrones de conducción con las vibraciones de la red atómica. A medida que la temperatura se reduce, la vibración de la red disminuye. En consecuencia, la conductividad de un metal aumenta a medida que la temperatura decrece. En última instancia, la conductividad viene limitada por la difusión de electrones en las im-

perfecciones de la red cristalina, en los átomos de impurezas y en otras fuentes tales como la superficie del cristal.

Los semiconductores y los aislantes difieren de los metales en que cada banda de energía está totalmente llena o totalmente vacía. En un semiconductor, la conducción es posible cuando se excita un electrón por medio de la luz o el calor y dicho electrón se mueve desde una banda llena a alguno de los niveles vacíos que existen en alguna de las bandas no ocupadas. El correspondiente cambio en la energía del electrón crea de hecho dos portadores de cargas: el propio electrón y el "agujero" que deja tras él en la banda previamente llena. La energía mínima necesaria para crear estos portadores de cargas es igual a la anchura del intervalo que separa a las dos bandas en cuestión. Los portadores de cargas se pueden suministrar también por medio de defectos o impurezas, tales como los que se utilizan en la preparación de los dispositivos semiconductores de silicio.

La conductividad de un semiconductor resulta determinada por dos factores: la concentración de los portadores de cargas y su movilidad o medida del tiempo medio durante el que se mueven dichos portadores sin sufrir ninguna dispersión. Debido a la excitación térmica de los electrones a través de la separación entre las bandas, la concentración de los portadores de carga aumenta con la temperatura, pero también lo hace la dispersión por las vibraciones de la red cristalina y, en consecuencia, la movilidad de los portadores de carga disminuye. En la mayoría de los semiconductores a temperaturas ordinarias la disponibilidad de portadores de carga es el factor predominante de la conducción. Como resultado de este hecho la conductividad disminuye cuando lo hace la temperatura, en contraste con lo que sucede en los metales.

Un aislante, al igual que un semiconductor, es un material que sólo tiene bandas totalmente llenas y bandas totalmente vacías, pero el intervalo prohibido entre las bandas es en general mucho mayor que el que existe en los semiconductores, hasta el punto de que la creación de portadores de carga por excitación térmica es prácticamente imposible. La conductividad neta que se observa en dichos materiales viene determinada principalmente por los defectos e impurezas presentes en la correspondiente estructura; en general, la descripción de este fenómeno no se lleva a cabo a partir del modelo de bandas, sino mediante modelos teóricos diferentes.

Todavía existe otra categoría de estructura de bandas. Ciertos materiales se describen formalmente como si sólo tuvieran bandas llenas y vacías, en una forma tal que las correspondientes energías se solaparan entre sí. En un sólido de este tipo, los electrones se redistribuyen para crear dos (o más) bandas parcialmente llenas, con el fin de que su energía sea mínima. Estos materiales se conocen con el nombre de semimetales y sus propiedades eléctricas no son sustancialmente diferentes de las de los metales; en particular, su conductividad aumenta a medida que la temperatura disminuye. El $(SN)_x$ es un ejemplo de semimetal.

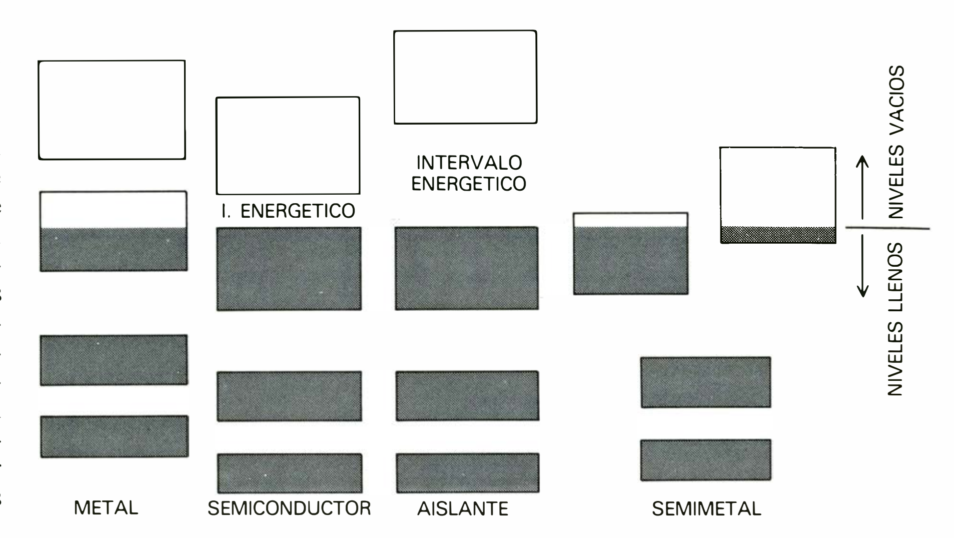
El solapamiento de unos orbitales determinados explica la formación de bandas anisótropas en materiales de cadena lineal; mas, para que se produzca una conductividad elevada, es necesario también que una de las bandas esté sólo parcialmente ocupada. Esta condición se puede cumplir incorporando al sólido grupos cargados cuyas componentes no guardan entre sí una relación entera sencilla; en tal caso se dice que el compuesto resultante no es estequiométrico. La fórmula dada anteriormente para la sal de bromuro potásico del TCP pone de manifiesto que en cada molécula existen 0,3 iones de bromo. Esto significa que, por cada 10 moléculas de TCP, se incorporan al cristal tres iones de bromo. Otras sales del TCP presentan un carácter no estequiométrico similar al de ésta.

En el (TTF) (TCNQ), que es un cristal estequiométrico, se transfiere la carga entre las dos columnas conductoras. El TTF es un dador de electrones y el TCNQ un aceptor de electrones; por término medio, desde cada molécula de TTF se transfieren 0,59 electrones a las pilas de TCNQ. Como resultado de este proceso se vacía parcialmente una banda llena y se llena parcialmente una banda vacía.

Ya se ha descrito con anterioridad el mecanismo de la conducción en el $(SN)_x$. El $(SN)_x$ es un semimetal en el

FORMACION DE BANDAS DE ENERGIA en un sólido. Es una consecuencia del solapamiento de los orbitales de átomos adyacentes. En un átomo aislado, cada uno de los orbitales tiene una energía bien definida y sólo puede alojar un máximo de dos electrones. Cuando dos átomos se acercan entre sí lo suficiente como para que sus orbitales se solapen, cada nivel energético se desdobla en dos; si añadimos un tercer átomo, los niveles atómicos se desdoblan en tres componentes. En un sólido constituido por unos 10^{23} átomos, los niveles son tan numerosos y están tan poco espaciados entre sí que, en realidad, forman una banda continua. La anchura de la banda depende de la cuantía en que los orbitales se solapan mutuamente. En un sólido constituido por *n* átomos, cada banda puede acomodar 2*n* electrones. Si cada uno de los átomos de un sólido posee solamente un electrón en un nivel energético determinado (en lugar del máximo de dos), la banda correspondiente únicamente estará medio llena.

que las bandas llenas y vacías se solapan en sus energías. El poliacetileno, cuando no está tratado, es un mal conductor debido a la escasez de portadores de carga. Este compuesto se convierte en un buen conductor agregándole impurezas que actúen como dadores o aceptores de electrones. Cuando se lleva a cabo un dopado intenso de este compuesto, se introduce un gran número de electrones en las bandas de conducción o de agujeros en la banda de valencia, lo cual confiere al material muchas de las propiedades que poseen los metales. Cuando el poliacetileno tiene una baja concentración de impurezas (menos del uno por ciento) se comporta en muchos aspectos como un semiconductor tradicional.



LA ESTRUCTURA DE BANDAS DE UN SOLIDO determina la mayoría de sus propiedades eléctricas. En una banda vacía no puede tener lugar conducción alguna debido a que, en dicha banda, no existen electrones. La conducción es también imposible en una banda totalmente llena, en razón de que la totalidad de la población de electrones en ella no puede tener un movimiento neto. Un metal es un material con una banda parcialmente llena en la que los electrones pueden adquirir una velocidad neta cuando algunos de ellos se desplazan hasta energías infinitesimalmente mayores dentro de la banda en que se encuentran. Un aislante sólo posee bandas totalmente llenas y bandas totalmente vacías, con una gran banda de energías prohibidas entre ambas precedentes. La estructura de bandas de un semiconductor se parece a la de un aislante, pero el primero de estos materiales posee algunos portadores de cargas móviles. Los portadores se pueden introducir en el material base mediante impurezas, defectos, carácter no estequiométrico o excitación de electrones desde la banda completa de energía más alta (la banda de valencia) o la banda vacía de energía más baja (la banda de conducción). Lo que distingue a un semiconductor de un aislante es que la separación entre dichas bandas es menor, en comparación, en un semiconductor. En un semimetal la banda de valencia se encuentra llena y la banda de conducción, por contra, vacía; pero las energías de estas bandas se solapan entre sí. Como consecuencia de este hecho, los electrones se redistribuyen entre ellas, dando lugar a dos bandas que están parcialmente llenas.

Una de las propiedades más significativas de los materiales de cadena lineal es la variación de su conductividad con la temperatura. Estos materiales se pueden clasificar en tres grandes categorías atendiendo a la diferente dependencia de su conductividad con la temperatura. A temperatura ambiente, la conductividad de los materiales a los que denominaremos como de la Clase I varía entre 10^{-6} mho/centímetro y 10 mho/centímetro; cuando la temperatura disminuye, su conductividad se reduce de forma regular; en algunos casos dicha disminución tiene lugar muy rápidamente. Un ejemplo de este grupo es la sal de cesio del TCNQ, cuya fórmula es $Cs_2(TCNQ)_3$.

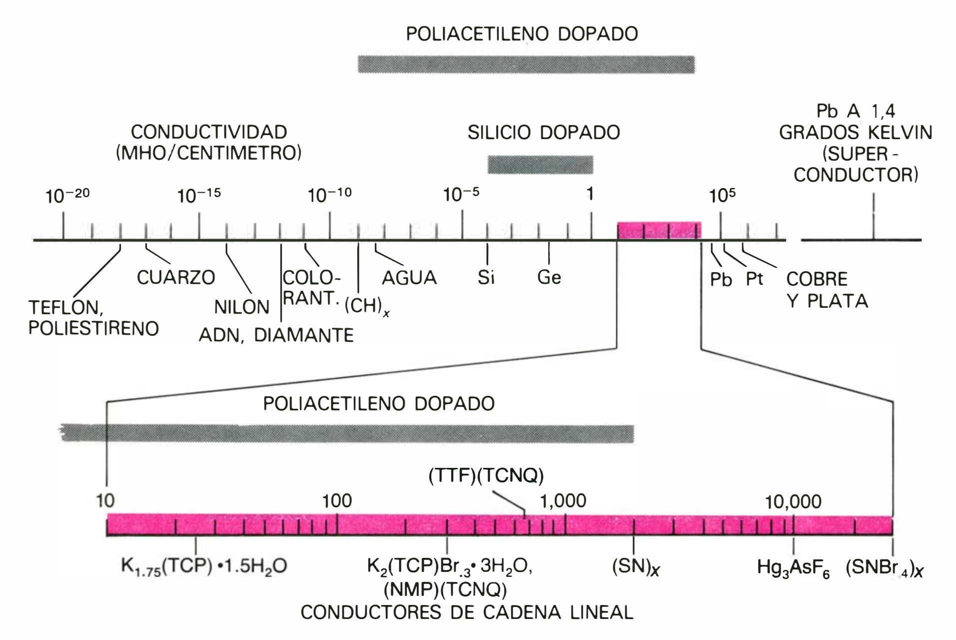
La conductividad de los materiales de la Clase II, medida también a temperatura ambiente, es del orden de 100 mho/centímetro. A medida que desciende la temperatura, la conductividad aumenta ligeramente hasta alcanzar un máximo muy amplio y, a partir de aquí, si la temperatura se sigue reduciendo, la conductividad disminuye fuertemente. Otra sal del TCNQ, formada en este caso con *N*-metilfenacinio, a la que nos referiremos como (NMP) (TCNQ), es típica de este grupo.

La Clase III está representada por el (TTF) (TCNQ). Su conductividad a temperatura ambiente es del orden de 650 mho/centímetro. A medida que la temperatura disminuye, la conductividad aumenta con regularidad, hasta que se alcanza un pico muy acusado de 10.000 mho/centímetro para una temperatura de 58 grados Kelvin. Si la temperatura sigue descendiendo por debajo de este valor, la conductividad disminuye muy rápidamente.

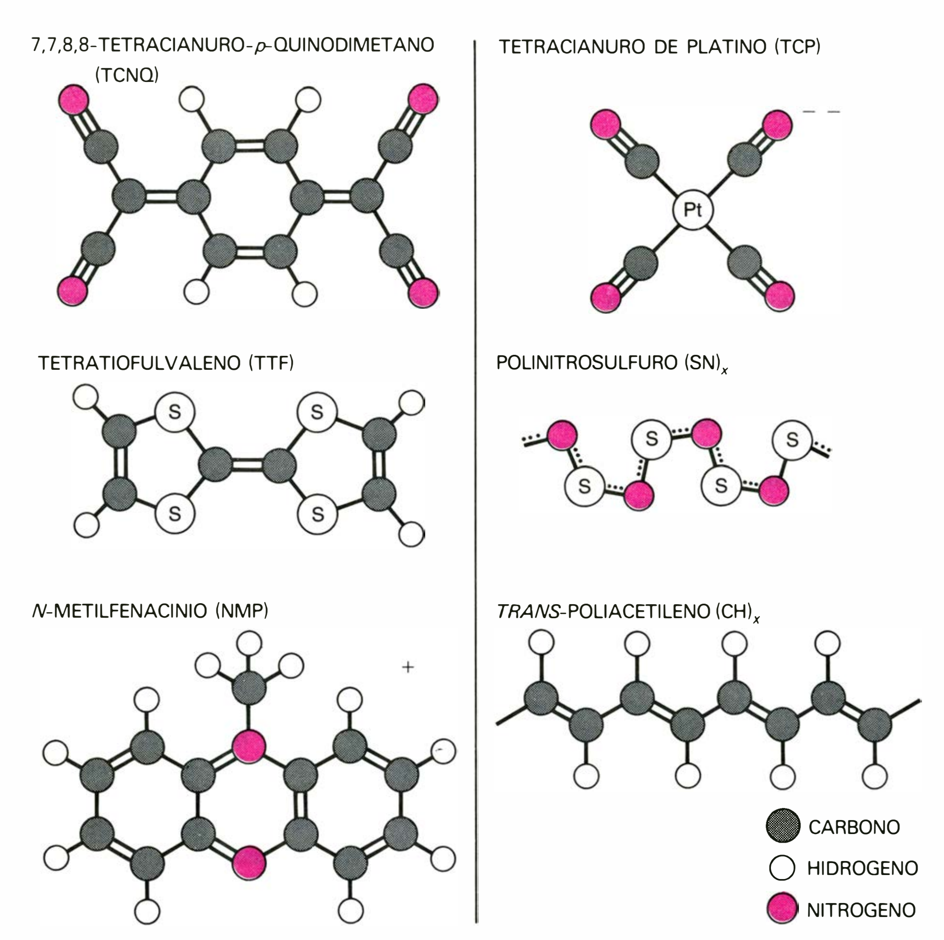
La dependencia de la conductividad con la temperatura para los materiales de la Clase I se puede interpretar con relativa facilidad. Estas sustancias se comportan como semiconductores en los que, para excitar los electrones o agujeros y, en consecuencia, activar los portadores de cargas, se necesita cierta energía térmica. La conductividad es menor a temperaturas bajas debido a que en tales condiciones se dispone de menos portadores de cargas. Los semiconductores no dopados, como el silicio, ofrecen un comportamiento similar.

Entre las diferentes aproximaciones propuestas para explicar el comportamiento de los sistemas de la Clase II, hay dos modelos teóricos que postulan la existencia de una transición de fase desde un estado metálico a altas temperaturas hasta un estado semiconductor a temperatura menor. Un modelo anterior estaba basado en una idea introducida en el año 1949 por Sir Nevill Mott, de la Universidad de Cambridge; en 1963 este modelo fue reformulado en términos más cuantitativos por John Hubbard, del Centro de Investigación de Energía Atómica (AERE) de Harwell, en Inglaterra. La hipótesis de trabajo de este modelo se basa en establecer los términos en que tiene lugar la ocupación de las bandas de energía de los sólidos. Aunque cada orbital puede acomodar un máximo de dos electrones de idéntica energía, Mott advirtió la existencia de una cierta repulsión electrostática entre dos electrones cualesquiera que se sitúen el uno junto al otro. Si la energía correspondiente a esta repulsión supera la anchura de la banda de energía en la que se mueven los electrones, dicha banda se desdoblará en dos partes. La mitad inferior se llenará cuando en todos los orbitales exista un electrón; evidentemente, una vez se haya cumplido esta condición, cualquier nuevo electrón que se añadiera al sólido tendría que alojarse en orbitales que ya estén ocupados por una de estas partículas. La adición de este segundo electrón a un orbital determinado exigiría una energía un tanto mayor que la normal, con vistas a vencer la oportuna repulsión. En otras palabras, se produciría una cierta separación entre las dos partes de la banda de energía a la que nos estamos refiriendo.

Cuando dicha banda se desdobla en dos partes, la mitad inferior está totalmente llena y la mitad superior está vacía, de forma que el material se comporta como un semiconductor y presenta la típica disminución de la conductividad para temperaturas bajas. Sin embargo, por encima de cierta temperatura umbral, la repulsión entre los electrones en el mismo orbital se ve superada por su movimiento térmico y surgen de nuevo las dos mitades de la banda de conducción.



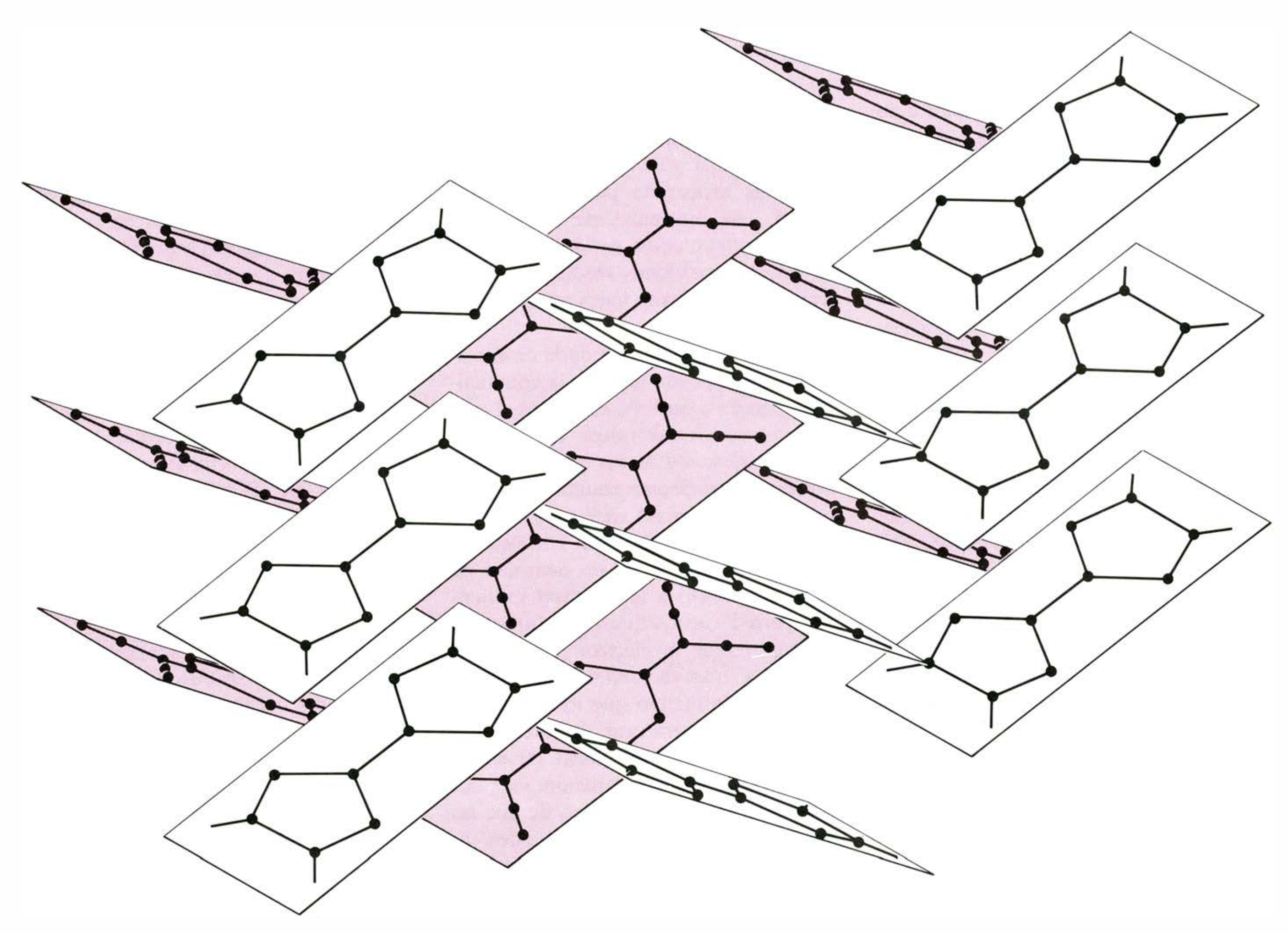
LA ESCALA DE CONDUCTIVIDADES cubre un intervalo enormemente amplio. La conductividad del cobre y la plata es 10^{24} veces mayor que la de aislantes tales como el teflón y el poliestireno. Las conductividades de la mayoría de los materiales de cadena lineal de alta conductividad varían entre 10 y 20.000 mho/centímetro. Los valores dados para cada material son típicos de ellos, pero existen grandes diferencias de una muestra a otra; el intervalo de los valores observados para el poliacetileno, un polímero orgánico, es muy grande. Todas las conductividades (si exceptuamos la del plomo superconductor) se hallan expresadas a temperatura ambiente; para los materiales de cadena lineal, las correspondientes medidas se han llevado a cabo en la dirección del eje de conducción que era más favorable.



SUBUNIDADES BASICAS de los conductores de cadena lineal. Se disponen en forma de columnas o de cadenas lineales. La unidad de tetraciamino de platino (TCP) tiene forma de disco y consiste en un ion de platino rodeado por cuatro grupos cianuro (CN^-). Las moléculas que representamos abreviadamente como TCNQ, TTF y NMP son también estructuras planas. Estas moléculas orgánicas son los constituyentes de las sales (TTF) (TCNQ) y (NMP) (TCNQ), en las que los dos tipos de moléculas citados en cada caso se disponen en forma de pilas independientes. El polinitrosulfuro y el poiiacetileno constituyen polímeros cuyas subunidades tienen forma de cadenas extensas y paralelas entre sí.

En los primeros años de la década de 1970, se aplicó esta teoría al sistema (NMP) (TCNQ), suponiendo que cada molécula de TCNQ transportaba una carga negativa y cada molécula de NMP transporta una carga positiva. El máximo observado en la dependencia de la temperatura con la conductividad se atribuyó a una transición de Mott-Hubbard. Actualmente se sabe que entre el NMP y el TCNQ se transfiere menos de una unidad de carga y que la transición de Mott-Hubbard no constituye un modelo adecuado para este sólido. Sin embargo, resulta evidente la necesidad de estudiar más detenidamente el papel que desempeña la repulsión entre los electrones en las propiedades eléctricas, magnéticas y ópticas de los conductores de cadena lineal.

En el año 1953, R.E. Peierls, de la Universidad de Birmingham, propuso otro tipo de transición de fase electrónica. Este autor señaló que todo conductor monodimensional podía sufrir una alteración del espaciado periódico de la red cristalina. Una hipótesis básica del modelo de bandas de los materiales monodimensionales consiste en la afirmación de que el espaciado entre las unidades de una columna o cadena es uniforme. Sin embargo, bajo ciertas condiciones, se puede reducir la energía de los electrones si se crea un intervalo prohibido en el nivel ocupado más alto de una banda de energías parcialmente llena. El resultado de este proceso es la conversión de un metal en un semiconductor, es decir, en un material que sólo posee bandas llenas y bandas vacías. Dicha banda prohibida se puede producir mediante una distorsión periódica de la red cristalina. En el caso más simple, en el que una banda está inicialmente semillena, las unidades de la cadena se agrupan formando pares, creando así unos espacios entre las unidades alternativamente estrechos y anchos. La distorsión resultante produce cierta separación entre el nivel ocupado de energía más alta y el nivel vacío de energía más baja, con lo cual el metal se convierte en un semiconductor. La distorsión aumenta la energía de la red cristalina. De aquí que la transición de Peierls tenga lugar sólo en el caso de que la energía de los electrones se reduzca lo suficiente como para compensar el aumento de la energía de deformación de la red cristalina. En un sólido cuasimonodimensional esta condición se cumple a bajas temperaturas; si la temperatura es elevada, la configuración preferente corresponde a una red cristalina regular. Se ha sugerido que en un derivado metilado del



FORMA DE ESPINA DE ARENQUE, constituida por los planos de las moléculas apiladas de la sal (TTF) (TCNQ). Dentro de cada pila, las moléculas se disponen paralelamente entre sí; mas, por otra parte, dichas moléculas están inclinadas respecto al eje de la pila. El solapamiento de los orbitales que se extienden por encima y por debajo del plano de cada molécula da lugar a una banda de conducción electrónica a lo largo de las pilas. Se ha encontrado que, por término medio, en cada molécula se transfieren 0,59 electrones desde el TTF hasta el TCNQ. Este proceso origina la aparición de bandas parcialmente llenas en ambas moléculas y, como resultado, se produce una conducción en ambos tipos de pilas. En el TCNQ los portadores de carga son electrones y en el tetratiofulvaleno (TTF) dichos portadores son agujeros (también llamados "huecos"): ausencia de electrones.

(TTF) (TCNQ) tiene lugar una transición de Peierls.

Diversos sistemas de la Clase II presentan un desorden cristalográfico significativo; algunos modelos hacen uso de este hecho para explicar la variación de la conductividad con la temperatura. El desorden produce un confinamiento de los electrones de un sistema monodimensional, de forma tal que dichos electrones no pueden moverse libremente a lo largo de las filas, sino que tienden a permanecer en una región cuyas dimensiones son relativamente pequeñas. En un estado monodimensional desordenado, la conducción resultaría limitada por la cuantía en que los electrones puedan saltar desde una región a otra. Este salto se vería facilitado por las vibraciones de la red y, en consecuencia, la conductividad resultaría ser mayor a temperaturas elevadas. Los modelos de este tipo han sido elaborados por científicos de los Estados Unidos (de la Universidad Johns Hopkins y la Universidad de Chicago) y de la Unión Soviética. Los investigadores de Brown-Boveri desarrollaron una idea similar. Estos autores propusieron representar un compuesto del TCP mediante un modelo en el que las filas conductoras están interrumpidas de vez en cuando por defectos perfectamente aislantes. Para que los electrones pudieran cruzar los intervalos de energía que separan a las bandas sería necesaria cierta activación térmica. Sin embargo, ninguno de estos modelos es capaz de dar cuenta sistemática de la forma exacta de la curva de la conductividad en función de la temperatura.

El esfuerzo realizado en la interpretación del mecanismo de la conducción en los materiales de la Clase III, y en especial en el (TTF) (TCNQ), ha cristalizado en un elevado número de modelos teóricos. Muchos físicos teóricos opinan que algunos de los datos experimentales no pueden ser explicados por un modelo de bandas ordinario, en el cual los portadores de cargas se mueven de forma independiente entre sí. En su lugar, proponen unos mecanismos de movimientos colectivos de cargas que guardan ciertas semejanzas con los que explican la superconductividad.

En los metales de tres dimensiones, la superconductividad se interpreta a la luz de una teoría formulada en 1956 por John Bardeen, Leon N. Cooper y J. Robert Schrieffer, de la Universidad de Illinois; en la actualidad, esta teoría se conoce con las siglas BCS. En este modelo la corriente sin resistencia no está basada en el movimiento de electrones individuales, sino de pares de electrones. La estabilidad de estos pares tiene su origen en una atracción indirecta: la carga negativa de un electrón atrae los iones positivos metálicos que lo rodean, creando así, a su alrededor, una región en la que la carga positiva resulta reforzada. Es esta carga precisamente la que atrae al segundo electrón. El fenómeno que acabamos de describir se observa tan sólo a

temperaturas extremadamente bajas, debido a que los pares se rompen con facilidad como consecuencia de la agitación térmica.

La clase de superconductividad a la que se refiere la hipótesis de Little depende también del apareamiento de electrones, pero en un contexto diferente. Los pares de electrones se mueven a lo largo de uno de los ejes de conducción y la interacción atractiva indirecta entre ellos se ve afectada por la presencia de los grupos próximos que se polarizan con facilidad.

Una posible explicación de la conductividad de los sistemas de la Clase III depende también de la interacción entre los electrones de conducción y la red cristalina que los rodea, pero no exige un apareamiento de los electrones. Aun cuando este modelo fue aplicado por primera vez al (TTF) (TCNQ) por Bardeen, el mecanismo en cuestión fue propuesto originalmente por Herbert Fröhlich, de la Universidad de Liverpool.

Como hemos visto, la transición de Peierls, que convierte un metal en un semiconductor, tiene lugar a través de una distorsión periódica de la red: la cadena se comprime en una región y se estira en otra. La distorsión tiende a impedir la conducción ordinaria del tipo de bandas, pero puede reforzar la conductividad a través de otro mecanismo. Los electrones de conducción producen una distorsión de la red monodimensional y, al mismo tiempo, se ven afectados por ella, tendiendo a concentrarse en aquellas regiones en las que la carga positiva es mayor, es decir, en los lugares donde la cadena está contraída. Una vez que los electrones se han agrupado de esta forma sobre una cadena individual, tales electrones pueden moverse conjuntamente a lo largo de las fibras en respuesta a la aplicación de un voltaje adecuado. Los grupos de electrones a los que nos acabamos de referir son transportados por la onda en movimiento de forma parecida a como una plancha de "surf" es arrastrada por una ola del océano. Las moléculas de la red sólo realizan movimientos periódicos alrededor de sus posiciones medias, pero los electrones se mueven hacia adelante a lo largo de las fibras.

El movimiento de una onda de densidad de carga puede reforzar la conductividad debido a que ello estabiliza el movimiento de los electrones. Un electrón que se mueva de forma independiente respecto a los demás resulta difundido con facilidad por un defecto o una vibración de la red cristalina. Un electrón atrapado en una onda de densidad de carga está sujeto a las mismas influencias, pero a menos que la perturbación sea importante, el electrón afectado es devuelto a su trayectoria original gracias al potencial atractivo que lo rodea.

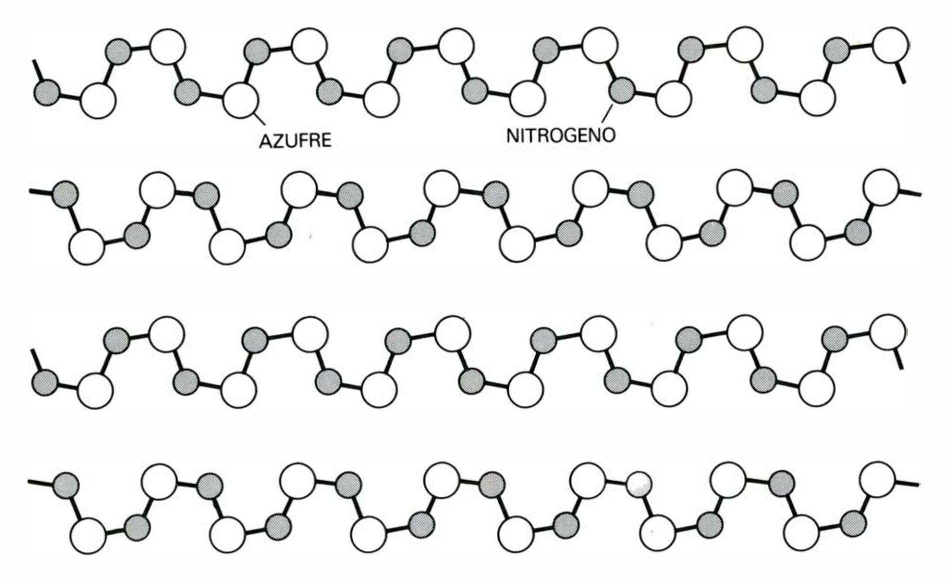
A bajas temperaturas, la interacción electrostática entre las ondas de densidad de carga vecinas aumenta; esta circunstancia da lugar al hecho de que las ondas anteriormente citadas queden inmovilizadas en el cristal. Del mismo modo, dichas ondas pueden quedar inmovilizadas sobre un defecto o cualquier otra irregularidad de la red cristalina; en consecuencia, la correspondiente conductividad dependerá de la calidad de una muestra determinada. La inmovilización de las ondas podría explicar la brusca disminución de la conductividad del (TTF) (TCNQ) para temperaturas inferiores a los 58 grados Kelvin. El modelo de Fröhlich se ha aplicado también al triseleniuro de niobio, que es un compuesto inorgánico cuasimonodimensional.

Para explicar la conductividad de los materiales de la Clase III sin recurrir al transporte colectivo de cargas se han propuesto otras muchas teorías. Por ejemplo, los investigadores de la International Business Machines Corporation han sugerido que las propiedades peculiares de estos materiales se pueden atribuir a la difusión de electrones por otros electrones. Un equipo de la Universidad Hebrea de Jerusalén ha formulado una explicación basada en la difusión de electrones que se produce en virtud de los movimientos vibracionales de las moléculas de TCNQ. Finalmente, un grupo de la Universidad de París ha señalado la posibilidad de que la difusión de electrones esté originada por excitaciones magnéticas.

La empresa Xerox ha desarrollado un modelo que se puede aplicar a los conductores orgánicos de cadena lineal que muestran los tres tipos de comportamiento conductor a los que nos hemos referido anteriormente. Para cumplir dicho fin es necesario limitar el intervalo de temperaturas a las situadas por encima de los 70 grados Kelvin. Por debajo de este valor, el transporte por electrones independientes entre sí no permite explicar los datos observados en todos los casos.

El modelo en cuestión se ocupa fundamentalmente de los factores básicos que determinan la conductividad de los semiconductores: la concentración de los portadores de carga, que se representa por el símbolo n, y la movilidad de dichos portadores, que se representa mediante el símbolo μ. La conductividad de dichos materiales viene dada por el producto de estas dos magnitudes y una constante, e, que representa la carga elemental del electrón, es decir, la conductividad es $ne\mu$. Sin embargo, la utilidad de esta expresión en la comprensión del fenómeno de la conductividad está condicionada por el hecho de que los valores de n y μ dependen de la naturaleza del material considerado y la temperatura. El modelo desarrollado por la Xerox analiza estas variaciones en los conductores moleculares orgánicos de cadena lineal.

En muchos casos, la variación de n con la temperatura puede representarse mediante una expresión exponencial en la que el exponente es el cociente entre dos energías. Una de estas energías es igual a la mitad del intervalo prohibido que separa el nivel ocupado más alto y la banda vacante de energía más baja; la

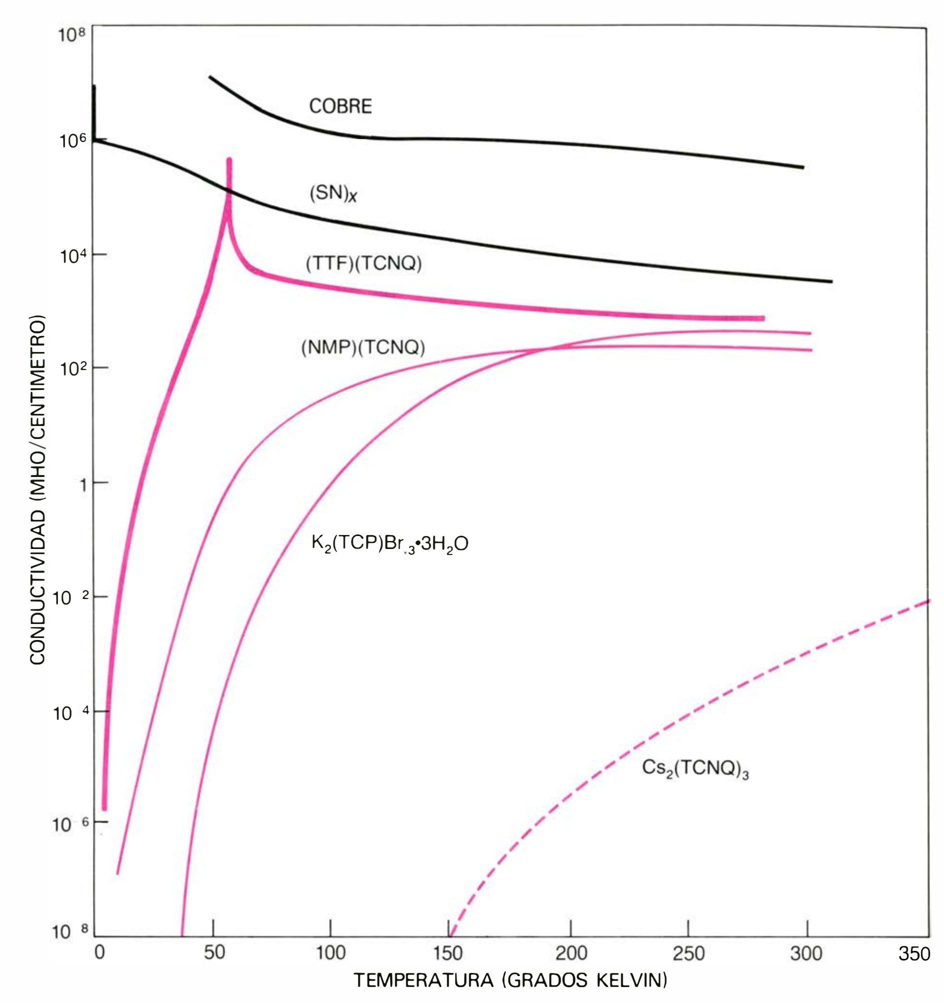


ESTE CONDUCTOR POLIMERO está constituido por átomos de azufre y nitrógeno alternados formando unas fibras largas que se alinean paralelamente entre sí. Las interacciones entre las fibras son significativas: la conductividad a lo largo del eje de la cadena es sólo unas 10 veces mayor que a lo largo de una dirección perpendicular a dicho eje. Este material, de fórmula $(SN)_x$ es un semimetal.

otra es la energía térmica, es decir, el valor de la temperatura expresado en unidades térmicas. La forma de esta expresión supone que cualquier disminución en el valor del citado intervalo prohibido o cualquier aumento en la temperatura dará lugar a que la concentración de los portadores de carga aumente exponencialmente. Esta relación está de acuerdo con la intuición, puesto que los portadores de carga se crean excitando térmicamente a los electrones para que crucen el intervalo prohibido que separa a las bandas.

Se ha comprobado que la movilidad es proporcional a la temperatura absoluta elevada a la potencia $-\alpha$, donde el exponente α es una constante cuyo valor depende de la naturaleza del material y de su calidad. Cualquier aumento de la temperatura disminuye la movilidad, lo cual concuerda también con la intuición, puesto que la movilidad resulta limitada por la difusión de electrones que se produce en virtud de las vibraciones térmicas de la red cristalina. La constante α determina cuán rápidamente disminuye la movilidad al aumentar la temperatura.

La representación gráfica de estas funciones pone de manifiesto que n aumenta considerablemente al aumentar la temperatura, al tiempo que μ disminuye también de forma significativa. La variación de la conductividad con la temperatura viene dada por la curva que representa el producto de n y μ. Dicha curva presenta un máximo para una determinada temperatura finita.



DEPENDENCIA DE LA CONDUCTIVIDAD con la temperatura. Es un indicador importante de la estructura electrónica de los sólidos. En metales como el cobre y semimetales como el $(SN)_x$, la conductividad aumenta a medida que la temperatura disminuye. En general, los semiconductores presentan un comportamiento diferente: la conductividad disminuye a medida que lo hace la temperatura. Los conductores moleculares de cadena lineal se pueden clasificar en tres clases, de acuerdo con las diferentes formas en que la conductividad depende de la temperatura. Los materiales de la Clase I, tipificados por una sal de cesio del TCNQ, se parecen a los semiconductores ordinarios en el hecho de que su conductividad disminuye regularmente cuando se enfrían. La Clase II está constituida por materiales tales como el (NMP) (TCNQ) y las sales de potasio del TCP cuya conductividad presenta un máximo muy amplio. El (TTF) (TCNQ) es un material de la Clase III: si este cuerpo se enfría a partir de la temperatura ambiente, su conductividad aumenta hasta alcanzar un valor máximo a 58 grados Kelvin y, a partir de ese instante, si la temperatura sigue descendiendo, la conductividad irá decreciendo.

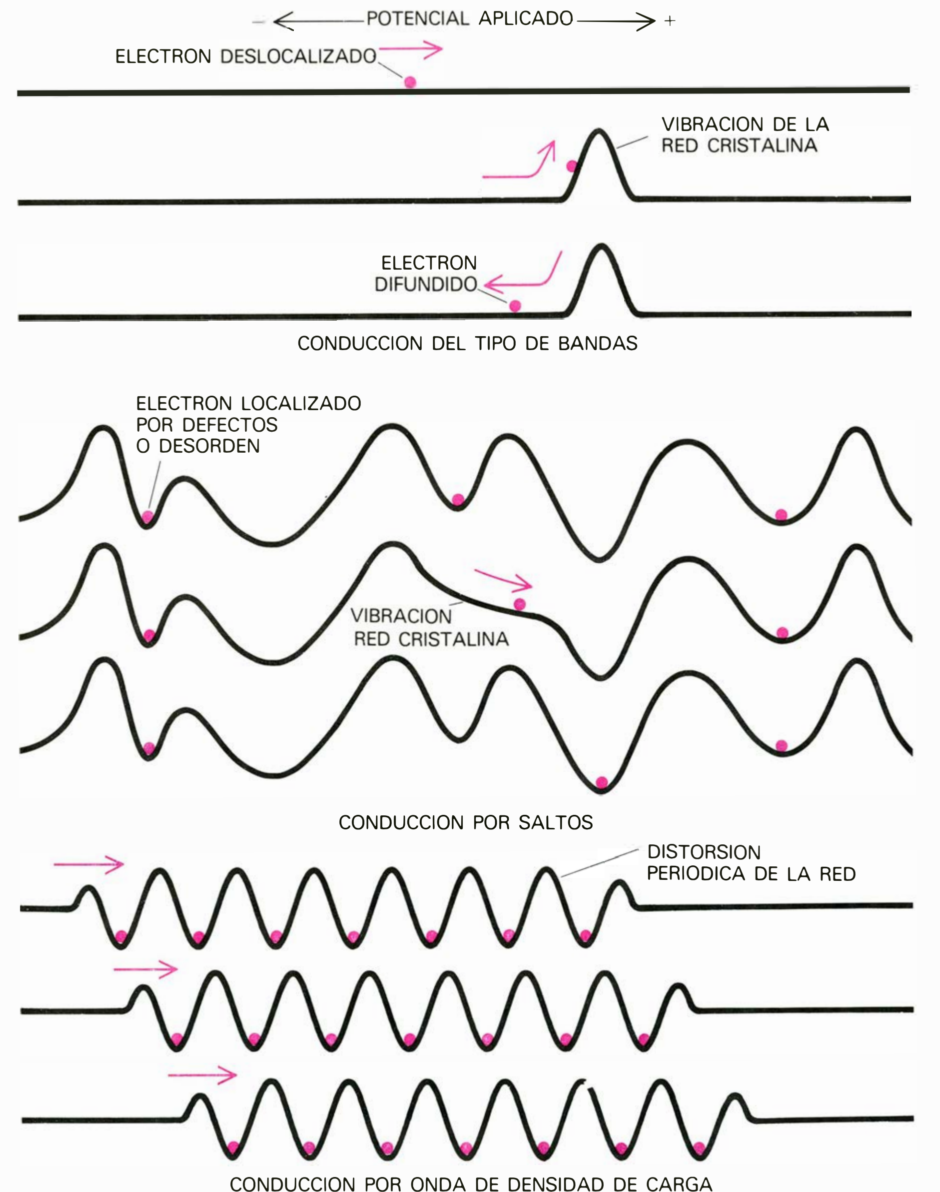
La forma de la curva teórica de la conductividad puede ponerse de manifiesto ajustando los datos experimentales correspondientes a los materiales de la Clase II. La conductividad de estos materiales presenta un máximo amplio para temperaturas un tanto inferiores a la temperatura ambiente. El modelo se puede aplicar también a los sistemas de la Clase I y la Clase III. La conductividad de los materiales de la Clase I es relativamente baja, aumentando con la temperatura. Este comportamiento puede interpretarse a la luz de dicho modelo, si se supone que el intervalo energético en estos materiales es sustancialmente mayor que para los sistemas de la Clase II. Al aumentar la separación entre las bandas, la concentración de los portadores de carga para una temperatura dada disminuye. El valor máximo de la conductividad se presenta para temperaturas tan elevadas que en la práctica son inalcanzables.

Se puede suponer que la estructura de bandas de los materiales de la Clase III presenta un intervalo energético despreciable. En consecuencia, la disponibilidad de portadores de carga se mantiene esencialmente constante, con independencia de la temperatura, y la conductividad viene gobernada principalmente por la movilidad. La movilidad es mayor a temperaturas bajas y, por tanto, la conductividad aumenta cuando el material se enfría, tal como sucede en los metales. Para temperaturas inferiores a los 70 grados Kelvin puede haber una contribución adicional a la conductividad del (TTF) (TCNQ) como resultado del movimiento de las ondas de densidad de carga. Por debajo de los 60 grados Kelvin, el material se convierte en semiconductor en virtud de una transición de Peierls.

Este modelo simplifica considerablemente la clasificación de los sistemas de cadena lineal. Lo que se había tomado como tres tipos de comportamiento diferentes se puede interpretar ahora como el resultado de una variación continua de un único parámetro: la anchura de la banda de energías prohibidas. El modelo que nos ocupa ha sido aplicado con éxito a la interpretación de otras propiedades físicas de los materiales de la Clase II, tales como las propiedades ópticas y la potencia termoeléctrica. Por otro lado, la movilidad de los portadores de carga de las sustancias de la Clase II y de la Clase III ha sido calculada directamente en la Xerox a partir de las constantes que definen el acoplamiento entre los electrones y las vibraciones de la red cristalina. Los resultados encontrados son coherentes con el modelo.

Con el fin de comprobar la aplicabilidad de los diferentes modelos teóricos propuestos a los materiales de la Clase

II, la Xerox desarrolló un nuevo sistema en el que la ocupación de las bandas de energía y el desorden del cristal pudieran variar de forma continua. El sistema es una variante del (NMP) (TCNQ) en el que una cierta proporción de las moléculas del NMP (hasta un 50 por ciento) se reemplaza por moléculas de fenacina. La fenacina tiene el mismo tamaño y forma que el NMP, pero permanece eléctricamente neutra a diferencia del NMP, que toma una carga. En consecuencia, por cada molécula de NMP reemplazada, el sistema (NMP) (TCNQ) pierde un electrón. La mezcla al azar de los dos tipos de moléculas aumenta también el desorden del cristal. La realización de una extensa serie de mediciones puso de manifiesto un comportamiento coherente con el de los materiales de la Clase II para todas las concentraciones de NMP estudiadas. La sustancia no podía convertirse en metálica cambiando la estequiometría, sino que se mantenía como un semiconductor en el que el intervalo prohibido era proporcional al cuadrado del contenido del NMP. Los estudios con rayos X llevados a cabo recientemente en Francia sugieren que el espaciado en la red cristalina no es uniforme y, por tanto, dicho intervalo puede ser del tipo Peierls. Estos resultados no están de acuerdo con las predicciones de los modelos para los que la disminución de la conductividad a bajas temperaturas es una consecuencia del desorden.

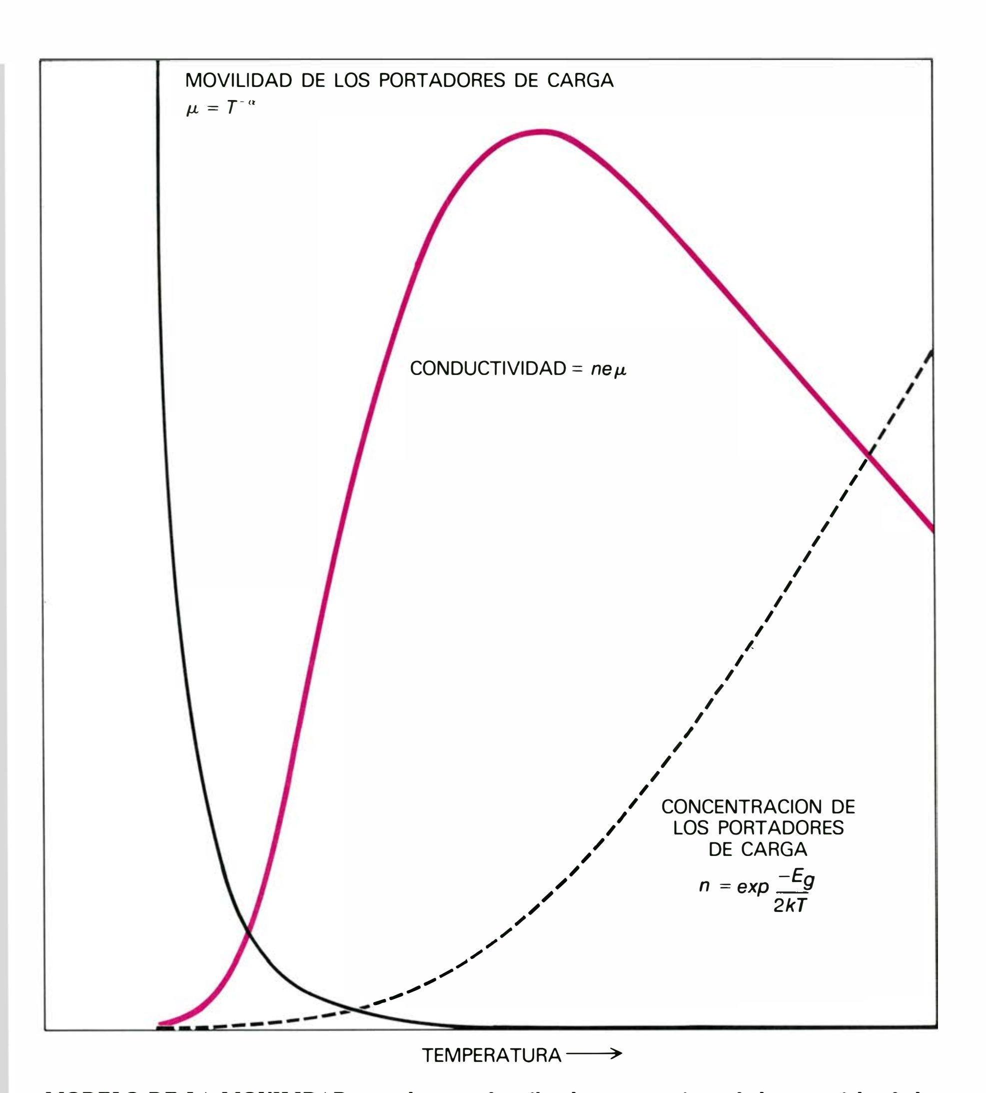


MECANISMOS DE LA CONDUCCION en los materiales cuasimonodimensionales; tales mecanismos comprenden algunos tipos que no se observan en los sólidos tridimensionales. En los metales (conducción de tipo bandas) la carga se transporta mediante electrones móviles y la conductividad se ve limitada por su difusión (o dispersión). Las fuentes de difusión más importantes son las vibraciones térmicas (representadas bajo la forma de un pico que el electrón en movimiento no puede remontar). Las vibraciones se hacen menos frecuentes a bajas temperaturas y, en consecuencia, la conductividad metálica aumenta al enfriar el correspondiente material. Algunos autores han sugerido que el paso de los electrones a través de las barreras de energía permite explicar las peculiaridades de la conducción anisotrópica. Dicha transición se ve favorecida por las vibraciones de la red, puesto que tales vibraciones pueden transferir la suficiente energía a un electrón localizado como para que éste sea capaz de superar la oportuna barrera y pueda alcanzar un lugar próximo al que ocupaba originalmente. El movimiento colectivo de los electrones se puede llevar a cabo mediante una onda de densidad de carga. Esta onda produce una distorsión de la red monodimensional en forma periódica, dando lugar a la creación de zonas cargadas positivamente en las que se acumulan los electrones. Cuando se aplica un potencial, los electrones se mueven arrastrados por la onda de densidad de carga anteriormente citada.

Los conductores de cadena lineal se han visto sometidos a un estudio intensivo durante estos últimos años y, al mismo tiempo, se ha progresado considerablemente en la comprensión de sus propiedades físicas. Estos progresos han contribuido a impulsar el trabajo de los químicos para preparar nuevos materiales. Por ejemplo, muchos cristales y polímeros de cadena lineal se han dopado recientemente con halógenos tales como el iodo y el bromo con el fin de aumentar su conductividad. Entre dichos materiales cabe citar el $(SN)_x$, el poliacetileno, las ftalocianinas y otros compuestos orgánicos.

En la Universidad McMaster de Ontario y en la Universidad de Pennsylvania se ha sintetizado un conductor anisótropo particularmente curioso. Se trata de un compuesto de mercurio, arsénico y fluor que responde a la fórmula Hg_3AsF_6. Este compuesto posee una estructura de cadena lineal en dos dimensiones. Las cadenas conductoras de los átomos de mercurio forman una serie de capas en las que todas las fibras que existen en una capa determinada están orientadas paralelamente entre sí. Sin embargo, en las capas inmediatas a la anterior, dichas fibras forman un ángulo de 90 grados respecto a las precedentes. Como resultado de esta estructura, la conductividad de dicho compuesto es elevada a lo largo de dos ejes perpendiculares entre sí, pero resulta ser menor en un factor de 100 a lo largo del tercer eje. Para temperaturas inferiores a los cuatro grados Kelvin, dicho compuesto se comporta como un superconductor.

A corto plazo, parece improbable que los materiales de cadena lineal cuya conductividad es elevada puedan reemplazar al cobre en su utilización como conductor eléctrico, pero no hay duda de que estos materiales encontrarán una serie de aplicaciones prácticas. El poten-



MODELO DE LA MOVILIDAD, con el que se describe el comportamiento de los materiales de las tres clases de conductores moleculares orgánicos de cadena lineal para temperaturas por encima de los 70 grados Kelvin. La conductividad viene definida por el producto de tres magnitudes: la carga del electrón (*e*), la concentración de los portadores de carga (*n*) y la movilidad de dichos portadores (μ). Dado que las vibraciones térmicas activan a los portadores de carga y los dispersan, la concentración de los portadores de carga aumenta con la temperatura pero su movilidad disminuye. El producto de estas curvas presenta un máximo para una combinación óptima de la concentración y la movilidad de los portadores. En los materiales de la Clase II dicho máximo se observa directamente. La separación entre las bandas de energía de los materiales de la Clase I es mayor que la de los anteriores y, como consecuencia, la temperatura para la que se predice que la conductividad será máxima está fuera de la región observable. Para los materiales de la Clase III dicha separación es despreciable y, por tanto, la movilidad gobierna a la conductividad; en virtud de esta relación, la conductividad aumenta al disminuir la temperatura. En las ecuaciones que definen las curvas, *T* representa a la temperatura en grados Kelvin, *kT* es la temperatura en unidades de energía, E_g la separación entre las bandas de energía (intervalo prohibido), α una constante determinada por los mecanismos de difusión del material y "*exp*" una notación que pone de manifiesto que la base de los logaritmos naturales tiene que elevarse a la potencia dada por el exponente que se indica en la correspondiente expresión. (George V. Kelvin.)

cial tecnológico del poliacetileno y otros polímeros similares a él constituye un caso particularmente digno de mención, debido a que las láminas dopadas de este polímero cubren un intervalo extremadamente amplio de conductividades y, al mismo tiempo, presentan todas las propiedades mecánicas que son típicas de estos polímeros, tales como la facultad de poder estirarse con facilidad.

Algunas de las aplicaciones de los materiales de cadena lineal se basan en el uso directo de las propiedades semiconductoras que son características de muchos de ellos. En la Universidad de Pennsylvania se han preparado láminas de polietileno dopadas con dos impurezas diferentes sobre un sólo diodo. Los investigadores de la Rockwell International Corporation han experimentado con una célula solar fabricada con $(SN)_x$. Finalmente, se ha comunicado al menos una aplicación comercial de estos materiales. En la Eastman Kodak Company se ha incrustado un conductor orgánico de cadena lineal en una lámina aislante con el fin de obtener una cinta fotoconductora para una máquina copiadora.

El microscopio acústico

Las investigaciones realizadas en la tecnología de la formación ultrasónica de imágenes han desarrollado un nuevo método experimental para obtener reproducciones con una resolución comparable con la del microscopio ordinario

Calvin F. Quate

Las ondas sonoras que vibran en la atmósfera con frecuencias inferiores a unos 20.000 hertz (oscilaciones por segundo) nos son familiares como medio de comunicación. Por ejemplo, las oscilaciones de la voz humana pertenecen a este grupo. La atmósfera no constituye un medio favorable para la propagación de ondas acústicas de frecuencias más elevadas, ultrasónicas. Tales vibraciones inaudibles se disipan rápidamente en gases tales como el aire. Las ondas de los ultrasonidos pueden propagarse a distancias apreciables con moderada atenuación sólo en líquidos y sólidos. Las ondas acústicas que vibran en tales medios condensados a frecuencias ultrasónicas se emplean para la formación de imágenes y no para la comunicación. Existen dispositivos ultrasónicos basados en la producción y detección de ondas acústicas con frecuencias del orden del megahertz (un millón de oscilaciones por segundo). Se emplean frecuentemente en el estudio de objetos submarinos, en averiguar características estructurales internas de materiales y de órganos internos del cuerpo humano. Por ejemplo, un feto en el seno de una mujer preñada puede ser examinado de este modo con completa inocuidad.

El microscopio acústico constituye una extensión de la tecnología de formación ultrasónica de imágenes. Las ondas ultrasónicas de tales microscopios tienen frecuencias próximas a un gigahertz (mil millones de oscilaciones por segundo), que son cerca de mil veces mayores que las frecuencias típicas de los sistemas macroscópicos de formación ultrasónica de imágenes. En términos de las longitudes de onda la comparación es ilustrativa. Se miden en metros las longitudes de onda de los sonidos producidos por la voz humana, en milímetros las de los ultrasonidos empleados en los dispositivos macroscópicos de formación de imágenes y en micrometros las de la mayoría de microscopios acústicos perfeccionados. Es decir, las ondas empleadas en estos nuevos dispositivos de formación de imágenes son de longitud comparable con la de las ondas electromagnéticas de la luz visible.

No hay que sorprenderse, por tanto, que el microscopio acústico empiece a proporcionar imágenes que son de resolución comparable con la de las imágenes conseguidas con el microscopio óptico ordinario. Lo que puede resultar extraño es que los que trabajan en este campo se están esforzando no en igualar la resolución del microscopio óptico sino en rebasarla. Pero no se trata solamente de la resolución. El origen del contraste en los sistemas acústicos es completamente distinto del de los ópticos. Ciertas propiedades, hasta ahora inaccesibles, de los objetos microscópicos empiezan a entrar en consideración en las imágenes acústicas. Se confía en que el microscopio acústico encontrará su lugar al lado del microscopio óptico, el microscopio electrónico y el futuro microscopio de rayos X, formando una línea de dispositivos complementarios para explorar el mundo de las cosas muy pequeñas.

Hace ya más de 40 años que S. Y. Sokolov, de la Unión Soviética, tuvo la idea de utilizar ondas acústicas del orden de los gigahertz en un sistema de microscopía. Pero la tecnología necesaria para producir tales ondas no estuvo disponible hasta comienzos de los años 1960. Gran parte de los primeros trabajos sobre dispositivos ultrasónicos de muy alta frecuencia fueron realizados por Hans E. Bömmel y Klaus Dransfeld en los Laboratorios Bell. Ha transcurrido poco más de una década desde que James S. Imai e Isadore Rudnick, de la Universidad de California en Los Angeles, se las ingeniaron para producir en helio líquido ondas acústicas de una longitud de onda de sólo 0,2 micrometros.

En fecha reciente, varios grupos, con técnicas algo diferentes, han emprendido la tarea de intentar desarrollar un sistema práctico de microscopía acústica. En las primeras fases del esfuerzo tuvieron un papel destacado los investigadores de la Zenith Radio Corporation, dirigidos por Adrianus Korpel y Lawrence W. Kessler. (Este último ha pasado a la Sonoscan, Inc., que fabrica una versión de rastreo con láser del microscopio acústico.) Nuestro grupo de la Universidad de Stanford ha perseguido el objetivo activamente durante cinco años. Hemos conseguido reducir la longitud de onda operativa en nuestro dispositivo de cinco micrometros a medio micrometro. Pero esto no es definitivo. Creemos que los instrumentos del futuro operarán con longitudes de onda más cortas que las conseguidas por Imai y Rudnick.

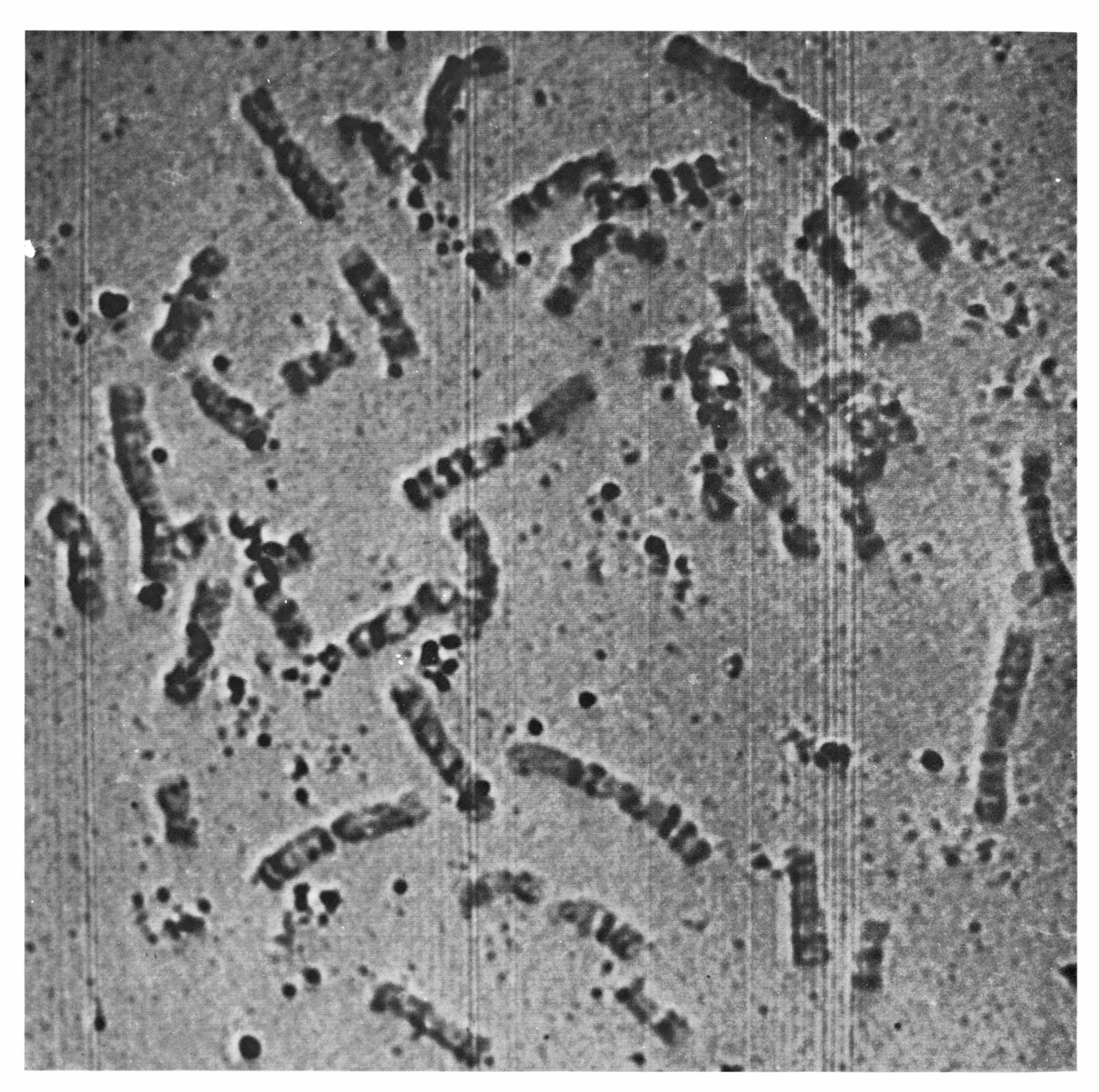
Mientras tanto, ¿qué es lo que hemos descubierto? Hemos aprendido que un microscopio acústico puede construirse con un elemento focalizador más simple que el constituido por las lentes del microscopio óptico ordinario. Hemos encontrado que la resolución de un microscopio acústico sólo está limitada por la longitud de onda con la que opera y que las aberraciones no desempeñan un papel importante. Hemos demostrado que los detalles en las micrografías acústicas pueden ser tan finos como en las micrografías ópticas. Hemos confirmado que es fácil medir el espesor de películas de semiconductores y metales con las ondas acústicas y que, por tanto, es posible estudiar la adhesión de tales películas a un substrato. En el caso de muestras biológicas, hemos conseguido imágenes con un elevado contraste de varios tejidos sin necesidad de una tinción previa. Confiamos en que esta cualidad del microscopio acústico permitirá que pronto pueda emplearse para lograr nueva información de las células vivas.

¿Cómo se forman las imágenes en un microscopio acústico? Al igual que en el microscopio óptico, la base reside en el cambio de la velocidad de propagación de las ondas al atravesar la superficie de separación existente entre dos medios

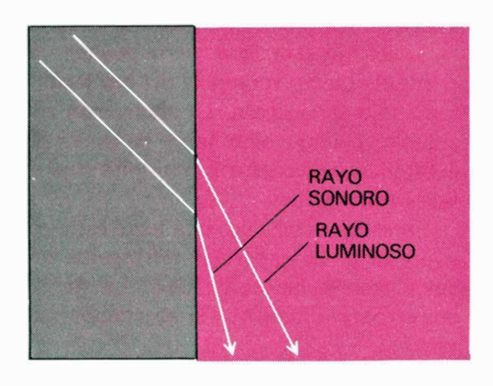
materiales distintos. En las ondas acústicas este cambio es mucho mayor que el que presentan las ondas luminosas. El índice de refracción, que mide las velocidades relativas de la luz entre dos medios, nunca es mayor de 1,9 y usualmente no vale más de 1,5. Por el contrario, la velocidad de las ondas sonoras puede cambiar en un factor diez al atravesar una interfacies (superficie de separación) sólido-líquido adecuada.

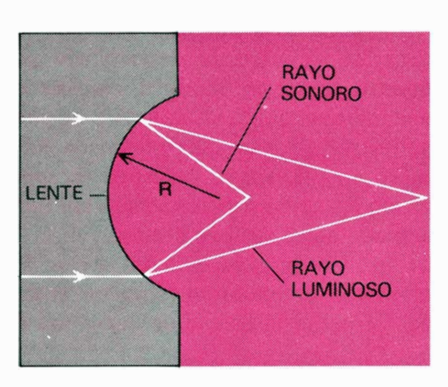
El elemento focalizador primario del microscopio acústico, desarrollado por nuestro grupo en Stanford, consiste en una interfacies esférica cóncava, de unos 40 micrometros de radio, formada entre zafiro y agua. En tal superficie de separación los rayos sonoros se desvían fuertemente hacia dentro al entrar en el líquido. En nuestro sistema todos los rayos convergen en un estrecho "embudo" próximo al centro de curvatura de la esfera. En dicha interfacies el cambio en la velocidad de la luz sería pequeño y el ángulo de convergencia del haz localizado sólo sería ligero. La aberración resultante debería eliminarse recurriendo a sistemas de lentes con múltiples superficies refringentes.

En el caso acústico, una lente esférica con una única superficie refringente resulta casi la hipótesis ideal. La fuerte convergencia hace que la aberración sea

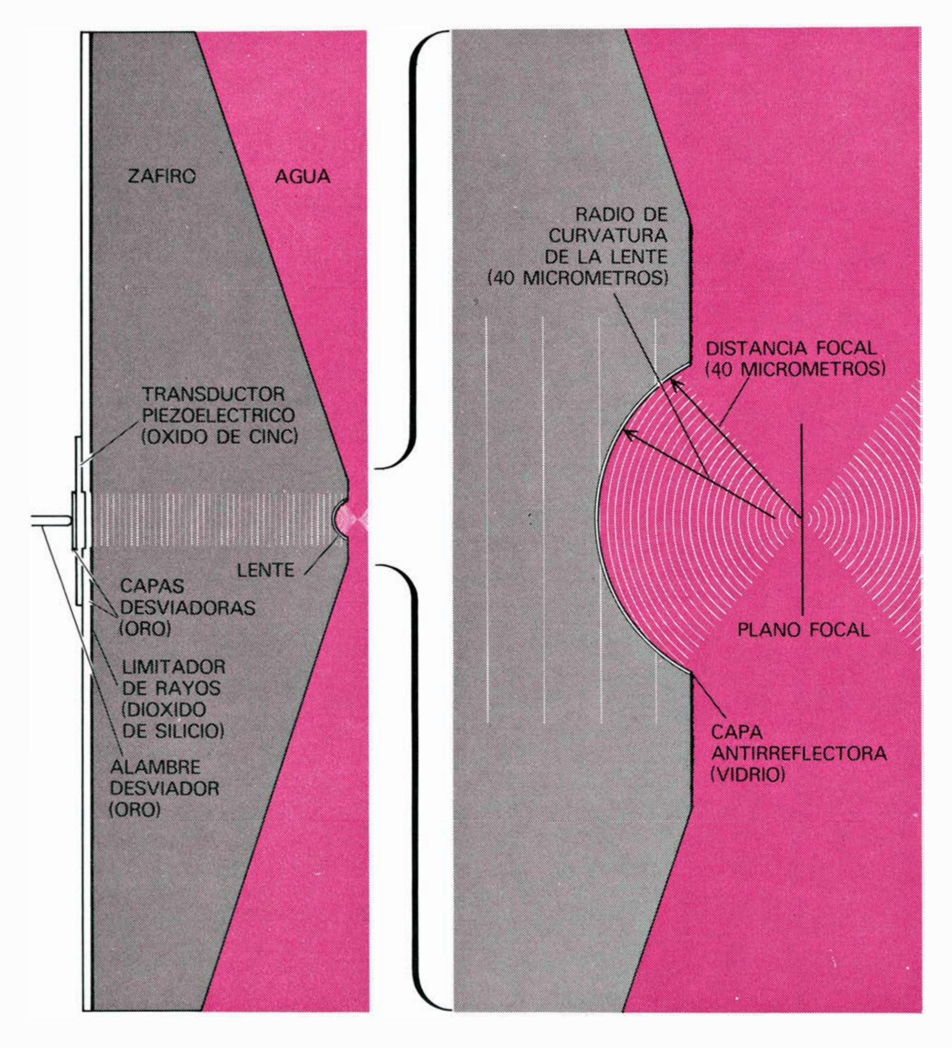


MICROGRAFIA ACUSTICA de un conjunto de cromosomas humanos, obtenida por el autor y sus colaboradores, en el departamento de Física aplicada de la Universidad de Stanford, como prueba del grado de resolución alcanzado por el nuevo dispositivo ultrasónico de formación de imágenes. Con este fin, tanto el portaobjetos con la muestra como la lente del instrumento estaban sumergidos en argón líquido. Se mantenían a una temperatura de 85 grados Kelvin (menos 188 grados Celsius) rodeando el recipiente que contenía el argón líquido de un baño con nitrógeno líquido. Se analizaron electrónicamente las señales ultrasonoras registradas para obtener la imagen de los cromosomas en una pantalla de televisor. Las líneas verticales que aparecen en la fotografía obedecen a defectos en el proceso de rastreo. El frotis de cromosomas fue preparado en el laboratorio de Howard M. Cann en la Facultad de Medicina de la Universidad de Stanford. Se eliminó el recubrimiento de proteína de los cromosomas y fueron teñidos de la forma habitual para poner de manifiesto en las micrografías ópticas su estructura en bandas. El aumento es de 3500 diámetros.





TANTO LOS RAYOS SONOROS COMO LOS RAYOS LUMINOSOS se refractan, o desvían, al atravesar la superficie de separación entre materiales en los que las velocidades de propagación de las ondas son distintas. El ángulo de desviación es mucho mayor en los rayos sonoros (*diagrama de la izquierda*). La presencia de una concavidad esférica en el material de mayor índice de refracción (esto es, en el que las ondas se propagan más deprisa) servirá para focalizar ambas clases de rayos (*diagrama de la derecha*). En esta situación, los rayos sonoros convergen más rápidamente que los rayos luminosos y en consecuencia la aberración resultante en la imagen obtenida es menor. Para las ondas sonoras, la distancia focal es 1,3 veces mayor que el radio de curvatura de la lente (*simbolizado con la letra R*). Para los rayos luminosos, la distancia focal es unas tres veces mayor que el radio de curvatura de la lente.



LENTE ACUSTICA diseñada por el grupo del autor en Stanford. Consiste en una superficie cóncava y esférica de separación entre zafiro y agua formada en la punta biselada de una corta varilla de unos 6 milímetros de diámetro (*izquierda*). La película piezoeléctrica que se encuentra depositada en la superficie posterior de la lente de zafiro convierte las señales eléctricas en otras acústicas durante el proceso de formación de la imagen. La misma película sirve a continuación para convertir las señales acústicas en señales eléctricas en la fase siguiente de lectura. La parte ampliada de la derecha muestra el perfil aproximado del haz de ultrasonidos focalizado en el agua. Las líneas blancas paralelas en ambos esquemas ilustrativos son los frentes de onda acústicos. La superficie lisa de la varilla de zafiro se ha hecho rugosa para dispersar las ondas espúreas. Además, se aplica usualmente un elemento antirreflector de vidrio a la superficie frontal de la lente al objeto de reducir las reflexiones acústicas internas.

mínima. Además, el recubrimiento antirreflector, semejante al comúnmente empleado en las lentes ópticas, puede aplicarse al zafiro para reducir la reflexión de las ondas acústicas en la interfacies. Sin embargo, las reflexiones residuales no pueden eliminarse completamente y constituyen una pequeña fuente de interferencia. Por fortuna, podemos distinguir las diferentes reflexiones con ayuda de una señal de prueba formada por un pulso momentáneo de ultrasonido. La parte del pulso que viaja a través del líquido se reflejará a partir del objeto y llegará al receptor más tarde que la parte reflejada por la lente. Este procedimiento permite identificar la causa de las ondas reflejadas, pues quedan registradas en distintos momentos.

Las principales componentes del sistema acústico son bastante simples, si se comparan con el prodigioso despliegue de equipo electrónico necesario para producir las señales y prepararlas para conseguir la imagen final. Se ha diseñado nuestro sistema de modo que podamos llevar las señales hacia dentro y hacia fuera de la celda acústica tan pronto como sea posible. La razón de esta manera de proceder reside simplemente en que, en la actualidad, las señales eléctricas son más fáciles de manipular que las acústicas.

El primer paso en el proceso de la obtención de la imagen consiste en convertir una señal eléctrica en otra acústica por medio de una película piezoeléctrica depositada en la superficie posterior de la lente de zafiro. La película está formada por numerosos cristalitos, cada uno de los cuales tiene una dirección preferida: su eje longitudinal, en general, es perpendicular a la superficie del substrato en el que está depositado. Un campo eléctrico alterno aplicado a lo largo de dicho eje comprimirá y expansionará el cristal según vaya variando la polaridad del campo. Desde hace años, se dispone de películas piezoeléctricas, las cuales forman ya parte de muchos dispositivos acústicos comerciales. La película piezoeléctrica de óxido de cinc que nosotros empleamos permite convertir la energía electromagnética en energía acústica o mecánica con un rendimiento que se aproxima al 50 por ciento.

La lente acústica focaliza el haz de ultrasonidos, producido de este modo, sobre una zona muy pequeña en el plano del objeto. La imagen se consigue moviendo dicha zona a lo largo del plano objeto, punto por punto y línea a línea, realizando un rastreo. El procedimiento

es análogo al de una cámara de televisión donde, luego de enfocar la imagen sobre una superficie fotosensible, se registra rastreando la superficie con un haz de electrones. Sin embargo, en el microscopio acústico el rastreo se realiza mecánicamente y su velocidad es mucho menor. Necesita varios segundos para completar el campo, lo cual hace necesario almacenar durante algún tiempo las señales registradas antes de que puedan ser utilizadas para formar una imagen del objeto.

Las señales reflejadas por el objeto y guardadas en la memoria electrónica terminan amplificándose y modulan la intensidad del haz de electrones en un receptor usual de televisión. La imagen se consigue rastreando el haz de electrones a través de la pantalla en sincronía con el movimiento realizado por el haz acústico al seguir el objeto (o, en nuestro modelo actual, con el realizado por el objeto respecto al punto focal del haz acústico). Se mantiene una correspondencia punto por punto entre las posiciones del impacto del haz electrónico sobre el monitor de televisión y la posición del impacto sobre el objeto del haz acústico. Si el impacto del haz electrónico se traslada un centímetro sobre la pantalla del televisor mientras el impacto del haz acústico se traslada sobre el objeto diez micrometros, la imagen resultante corresponderá a un aumento de mil diámetros. Puede ajustarse fácilmente la razón entre estas traslaciones y con ello la amplificación de las imágenes puede variar de cien a varios millares.

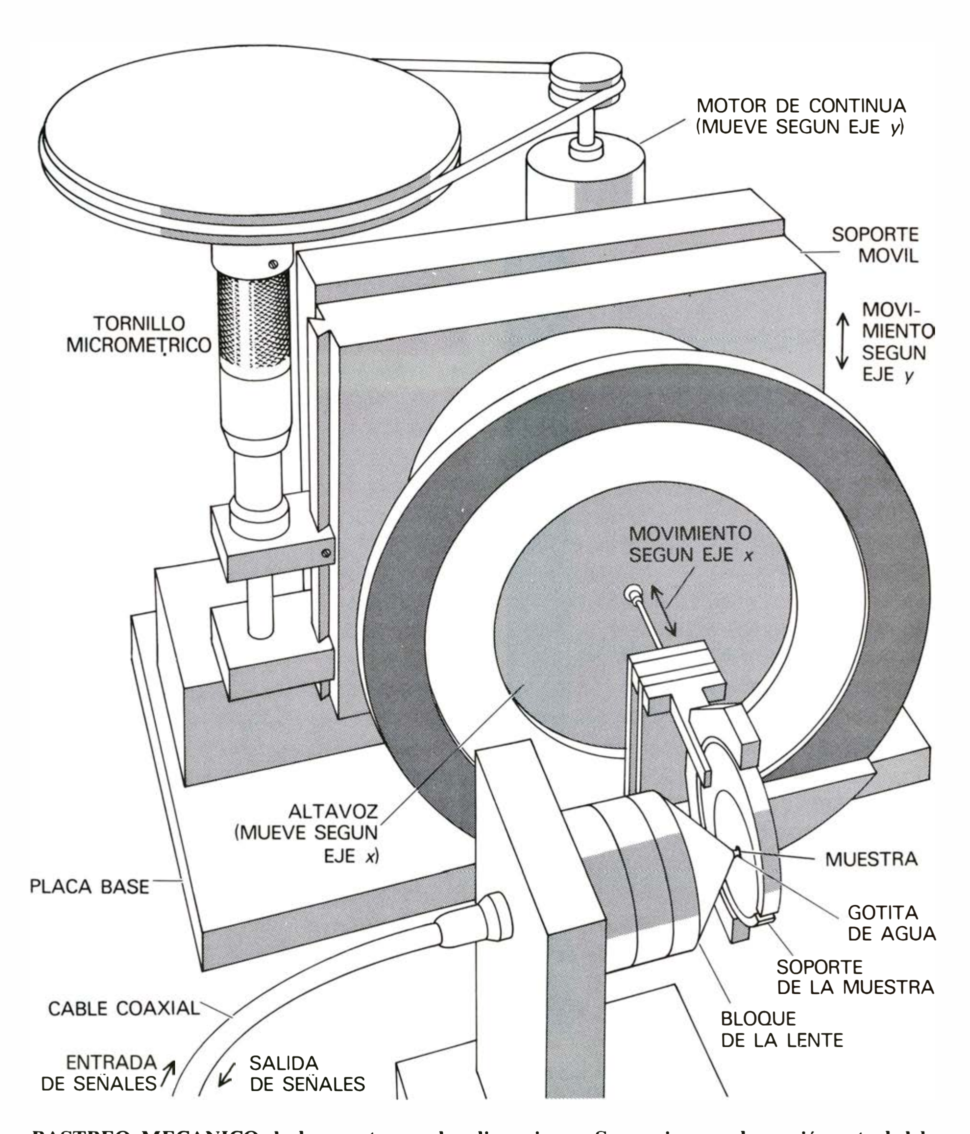
En nuestro sistema actual, el rastreo mecánico se consigue de la manera siguiente. Se fija la muestra a un disco circular unido a un soporte que es impulsado horizontalmente por un altavoz. El altavoz vibra con una frecuencia de 60 hertz y origina corrimientos horizontales de la muestra comprendidos entre 100 y 200 micrometros. El conjunto puede trasladarse en dirección vertical, pues el altavoz va montado sobre guías verticales. El movimiento vertical, impulsado por un motor de corriente continua, es mucho más lento que el movimiento horizontal.

Los visitantes que acuden a nuestro laboratorio suelen plantear dos cuestiones: 1.ª ¿No introduce, el movimiento relativo entre la muestra y la lente, vibraciones perturbadoras en la parte sólida del aparato o bien turbulencia en el líquido? 2.ª ¿Por qué no se sustituye el método de rastreo mecánico por otro electrónico, que eliminaría la necesidad de mover todos los componentes y así podría ser más rápido?

Respecto a la primera cuestión, algunas veces se presentan vibraciones en la muestra pero son muy pequeñas, probablemente debido a que los modos de resonancia mecánica del objeto son de frecuencias muy superiores a los 60 hertz del proceso de rastreo. En cuanto a la turbulencia en el líquido, hasta el momento no se ha observado. A medida que perfeccionemos el instrumento y observemos detalles más finos, estos factores podrán resultar importantes, pero no es ése el caso por ahora. Las imágenes conseguidas con el instrumento están bien enfocadas con bordes nítidos en todo el campo de visión.

La respuesta a la segunda pregunta va ligada a las posibilidades de invención. Apreciamos ciertamente las ventajas potenciales del rastreo electrónico, pero todavía no hemos logrado dar con un método no mecánico eficaz de rastreo con un haz bien enfocado.

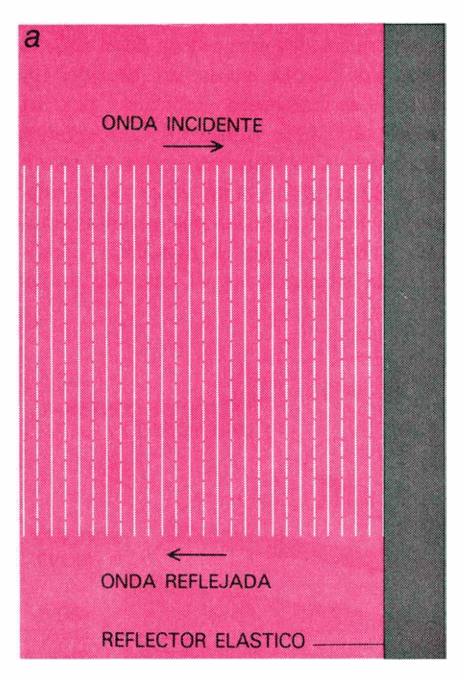
El examen de las superficies sólidas con este instrumento se realiza exponiéndolas a las ondas esféricas convergentes que proceden de la lente acústica y luego registrando las ondas divergentes reflejadas con ayuda de la misma lente que ahora actúa en sentido inverso. Las imágenes generadas por las ondas reflejadas muestran propiedades especiales. Conviene entender la razón física de las diferencias entre estas imágenes y las producidas por sistemas ópticos. Debido a que es más fácil imaginarse la situación en el caso de ondas

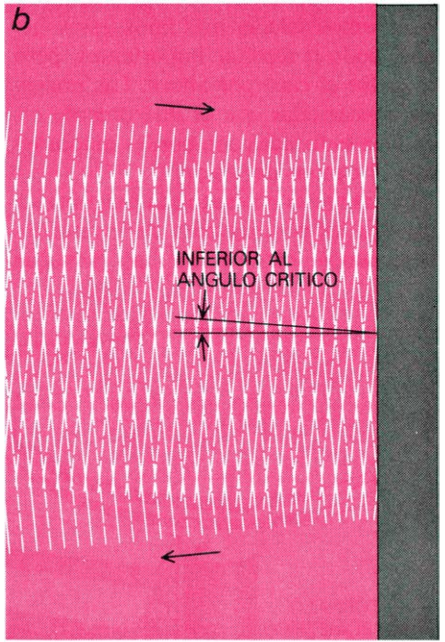


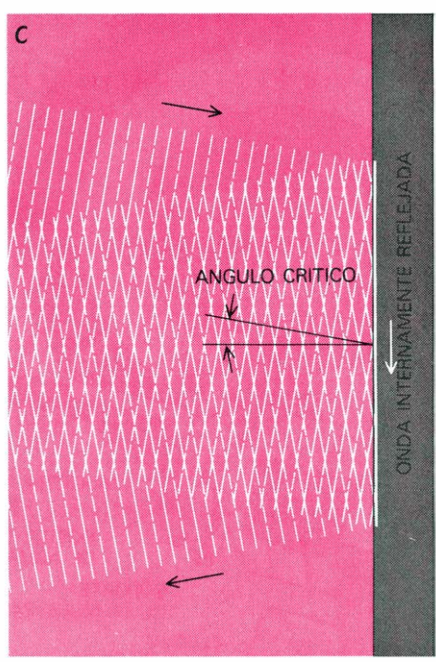
RASTREO MECANICO de la muestra en dos dimensiones. Se consigue en la versión actual del microscopio acústico de Stanford por medio del aparato representado. El disco circular que sostiene la muestra se desplaza horizontalmente entre 100 y 200 micrometros por un altavoz que vibra con una frecuencia de 60 hertz (oscilaciones por segundo). El corrimiento vertical del conjunto muestra-altavoz está gobernado por un motor eléctrico de corriente continua. El movimiento vertical es mucho más lento que el horizontal. Las señales acústicas reflejadas por la muestra se analizan electrónicamente y se destinan a modular la intensidad del haz electrónico de un televisor. La imagen se forma rastreando el haz electrónico sobre la pantalla sincrónicamente con el movimiento de la muestra relativo al punto focal del haz acústico. La razón entre corrimientos de los dos haces da el aumento de la imagen final.

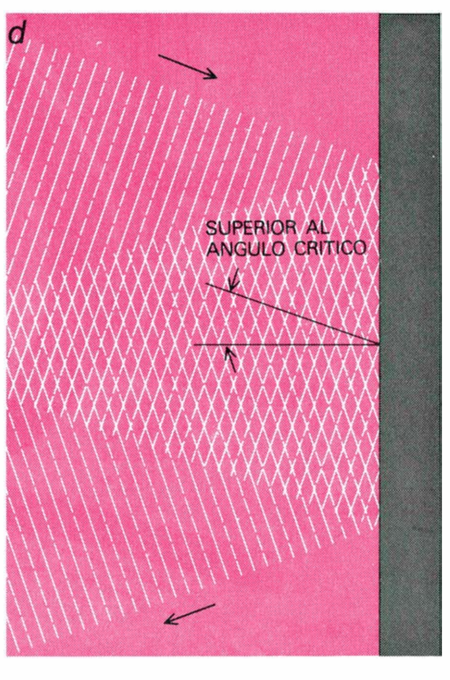
planas, describiremos las diferencias esenciales para éstas y después discutiremos el proceso general para ondas esféricas, convergentes y divergentes.

Al incidir una onda plana perpendicularmente sobre una superficie elástica plana se reflejará según la dirección opuesta con una fase idéntica a la de la onda incidente. Sin embargo, cuando la onda incidente avanza en una dirección inclinada respecto a la normal puede excitarse una nueva onda, que aparece pegada a la superficie. Si esto ocurre, la onda reflejada vuelve con un ángulo igual al de incidencia, pero con una fase que es distinta de la correspondiente a la onda incidente. Este fenómeno, denominado reflexión total interna, ha sido intensamente estudiado en las ondas luminosas que inciden en superficies transparentes por P. K. Tien, de los Laboratorios Bell, y otros autores. Forma la base de la nueva tecnología de la óptica de ondas guiadas [véase "Optica de ondas guiadas" por Amnon Yariv: INVESTIGACION Y CIENCIA; marzo, 1979].









ANGULO CRITICO para la reflexión interna. Es un concepto importante para comprender el fundamento físico de las propiedades especiales de las imágenes conseguidas en microscopía acústica. Puede hacerse visible en términos esquemáticos con ayuda de la ilustración, en la que tanto las ondas incidentes como las reflejadas se han representado como si fueran ondas planas. Cuando una de tales ondas incide perpendicularmente (*a*) será reflejada en dirección opuesta con una fase idéntica a la de la onda incidente. Los frente de las ondas planas se han representado alternando líneas blancas de trazo seguido y de trazo interrumpido. Los trazos seguidos designan crestas de las ondas, los interrumpidos valles. Cuando la dirección de la onda plana incidente se inclina respecto a la normal formando un cierto ángulo (*c*), a lo largo de la superficie del material reflector se excita una onda reflejada internamente y se altera la fase de la onda reflejada. (Se ha supuesto que el salto en la fase fue de media oscilación.) Si los ángulos de inclinación son menores o mayores que el crítico (*b, d*) las ondas se reflejan en fase. En microscopía acústica las ondas reflejadas internamente viajan por todas las superficies.

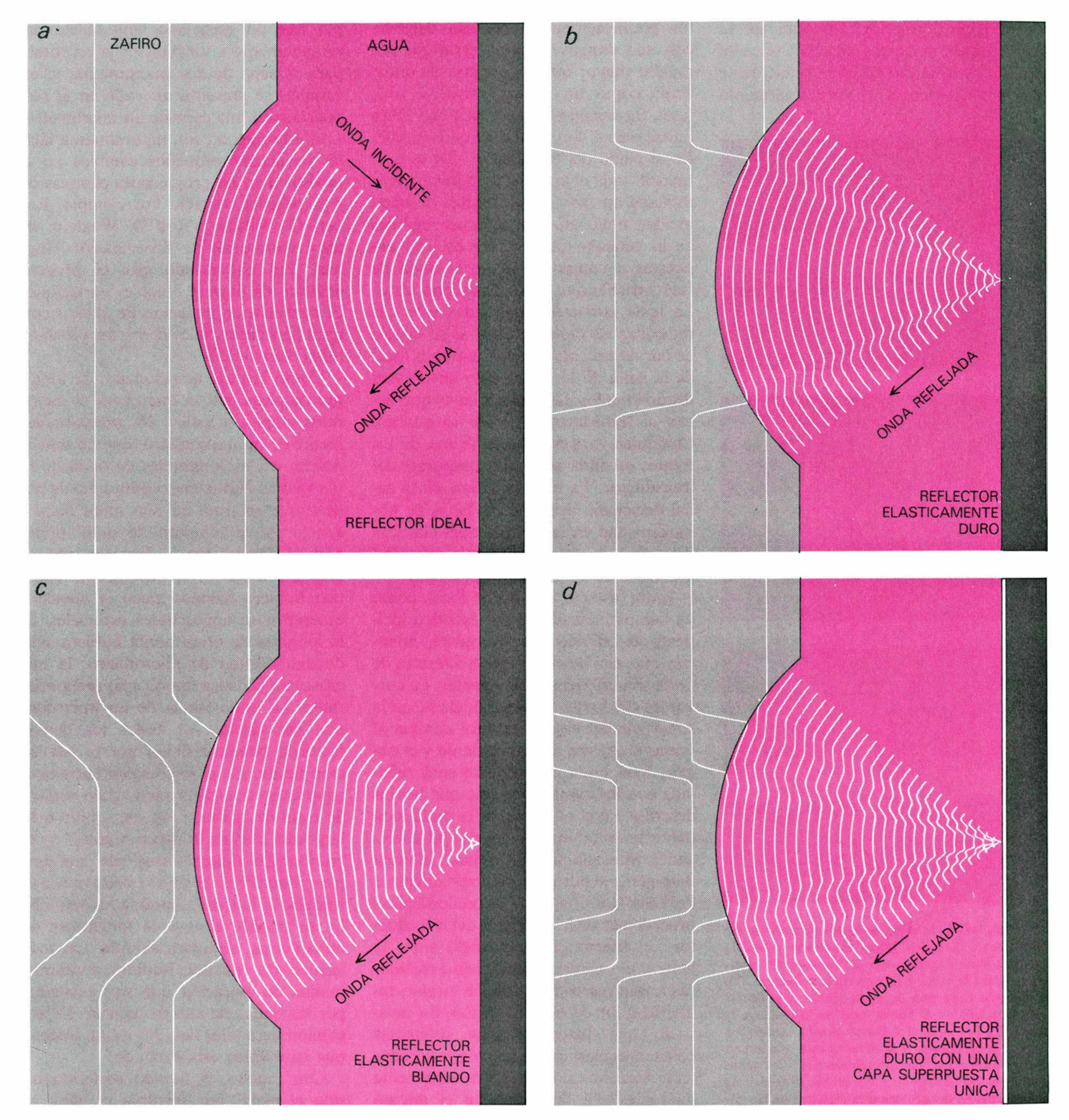
En los sistemas acústicos estas ondas especiales viajan a lo largo de las superficies libres y de las interfacies. Pueden ser excitadas ajustando la dirección de las ondas incidentes con el valor del ángulo crítico. Para ángulos menores o mayores que el crítico, la onda se refleja totalmente en fase con la incidente y no se transmite energía vibratoria hacia el interior del sólido. Para ondas sonoras incidentes en una superficie de separación líquido-sólido, el ángulo crítico puede llegar a ser de sólo diez grados.

K. W. Andrews y R. I. Keightly, de la British Steel Corporation, han empleado la reflexión total interna de ondas sonoras de alta frecuencia para examinar diversos materiales sólidos. Es sabido que las propiedades elásticas de los materiales cambian al aumentar los esfuerzos que soportan, y sería de esperar que dicho cambio afectase a la velocidad de las ondas sonoras que se propagan a lo largo de la superficie del material. Andrews y Keightly han demostrado que el ángulo crítico para la reflexión total interna varía, no sólo con el esfuerzo aplicado, sino también con la composición del material. En su aparato se determinan los ángulos críticos exponiendo la superficie a un haz de ondas sonoras planas, inclinado varios grados respecto a la posición de incidencia perpendicular. El transductor piezoeléctrico, para la recepción de las señales reflejadas, está también inclinado los mismos grados respecto a la perpendicular aunque en dirección opuesta. Con este sistema experimental pueden examinarse las faltas de homogeneidad en las propiedades elásticas de los materiales, que resultan, por ejemplo, de los cambios de composición en las aleaciones.

Como ya hemos dicho, en nuestro microscopio acústico las ondas ultrasonoras incidentes son esféricas y convergentes y el contraste que aparece en las imágenes procede de los complejos detalles del proceso de reflexión en ondas esféricas. En términos simples, puede decirse que los frentes de onda se reflejan en fase excepto en el sector de la onda próximo al ángulo crítico. A causa de la reflexión total interna, este sector

se refleja con una fase diferente. Las consecuencias de este sutil corrimiento de fase son decisivas.

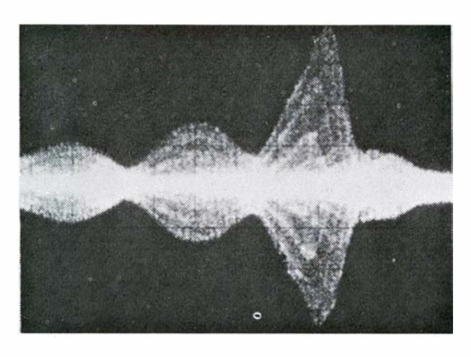
Consideremos los frentes de onda después de experimentar la reflexión. Supongamos que en la transición en fase cerca del ángulo crítico para reflectores elásticos los frentes de onda reflejados se trasladen toda una longitud de onda, es decir, de una cresta a la siguiente. Para un material duro, el zafiro, por ejemplo, la transición será brusca y los frentes de onda reflejados permanecerán aproximadamente esféricos. En un material blando la transición será gradual y los frentes de la onda reflejada resultarán distorsionados. Mostrarán poca relación con las ondas esféricas originales. En el caso de un substrato con una sola capa superpuesta hay dos interfacies reflectoras y dos ángulos críticos, el primero de-

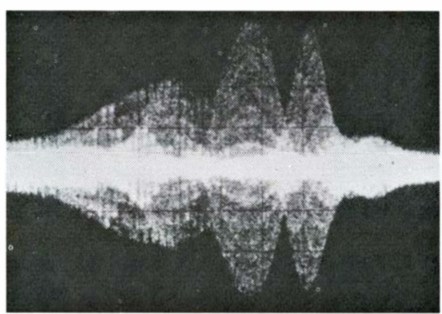


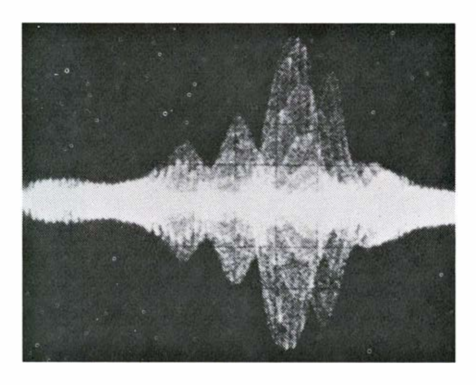
ONDAS ESFERICAS CONVERGENTES de ultrasonidos formadas por la lente zafiro-agua en el microscopio acústico. Se han representado después de ser reflejadas por cuatro superficies reflectoras distintas. En el caso de un reflector ideal (*a*) los frentes de onda reflejados divergentes coinciden exactamente con los frentes de onda que inciden convergiendo. Las otras tres superficies reflejan en fase las ondas incidentes, excepto en el sector de la onda próximo al ángulo crítico. Debido a la reflexión total interna, en aquel sector las ondas salen reflejadas con la fase opuesta. En el caso de un material elásticamente duro (*b*), la transición de ciclo completo en el ángulo crítico es abrupta y los frentes de onda reflejados permanecen aproximadamente esféricos. En un material elásticamente blando (*c*), la transición es gradual y los frentes de onda reflejados muestran distorsión. En un sustrato con una única capa superpuesta (*d*), la distorsión es aún mayor. Las ondas reflejadas se comportan de un modo más complicado que el expuesto.

terminado por el material de la capa superpuesta y el segundo influenciado por el material del substrato. Ya que los frentes de onda reflejada deben trasladarse dos longitudes de onda sucesivas, la distorsión es aún mayor que con una sola superficie blanda.

(Para mayor claridad, se ha simplificado la forma de onda en la ilustración de la página anterior. En realidad, la amplitud de la onda reflejada nunca es igual a la amplitud de la onda incidente y, en consecuencia, el comportamiento real de las ondas se vuelve más complicado. Aunque no hemos medido la forma real de las ondas, estamos seguros que este modelo ofrece los rasgos esenciales de la cuestión.)







AMPLITUD de la señal reflejada procedente del microscopio acústico. Está mostrada sobre la pantalla de un osciloscopio en función del espaciado entre la lente y la superficie reflectora. En el pico de cada una de estas trazas el reflector está colocado en el punto focal de la lente. A la izquierda del pico, el reflector queda expuesto a un haz convergente. A la derecha está expuesto a un haz divergente. Las diferencias de forma de las porciones oscilantes de estas trazas proceden de diferencias en las propiedades elásticas de las superficies reflectoras. La traza superior corresponde a una superficie de sílice pura. La central pertenece a una superficie de sílice cubierta con una capa de aluminio de un micrometro. La traza inferior está producida por una superficie de sílice con una capa de dos micrometros de aluminio. De este modo, pueden detectarse hasta cambios extraordinariamente pequeños en el espesor.

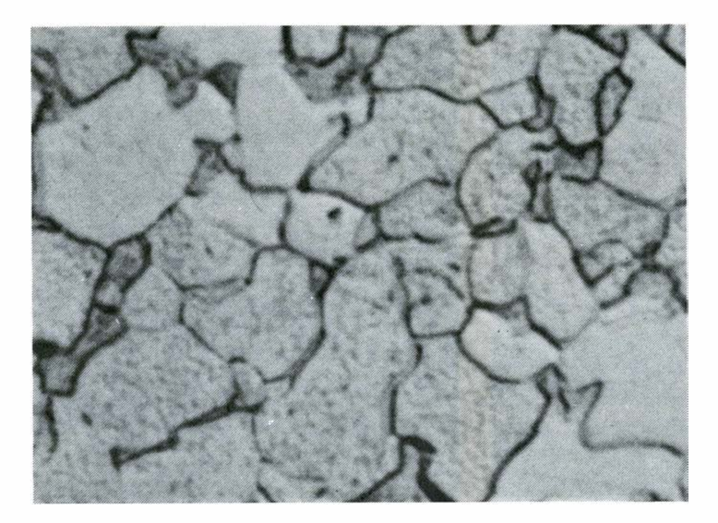
¿Cómo detectamos los frentes de onda distorsionados? Lo hace el transductor piezoeléctrico. Un transductor piezoeléctrico de película delgada, con una dimensión transversal que sea mucho mayor que la longitud de onda, actúa como un detector sensible a la fase. Su respuesta será máxima para ondas planas con frentes de onda paralelos al plano del transductor. En esta situación toda el área del transductor será excitada en fase. Si los frentes de onda inciden oblicuamente, algunas regiones de la película transductora estarán expuestas a componentes de la onda de fase variable que se anularán con otras. La lente convertirá las ondas esféricas reflejadas en ondas planas, si su centro de curvatura coincide con el punto focal de la lente. Si los frentes de onda están distorsionados, las ondas planas que llegan al transductor también lo estarán. Cualquier variación en la forma de los frentes de onda alterará la respuesta del transductor. Es esta variación en la señal detectada, debida a las faltas de homogeneidad en las propiedades elásticas del objeto, la que constituye la causa del contraste en las imágenes acústicas.

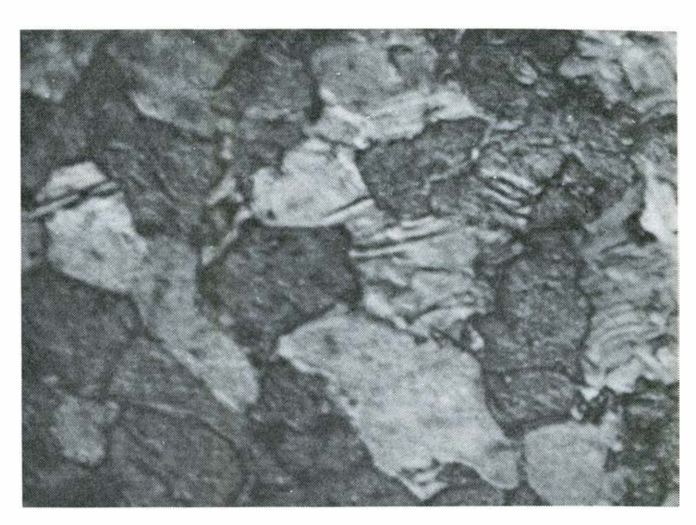
Entre la lente y el punto focal, como ya hemos indicado, el haz acústico incidente es esférico y convergente, mientras más allá de este punto los frentes de onda son esféricos y divergentes. La curvatura del frente de onda incidente en la superficie del objeto reflector cambia al variar el espaciado entre la lente y el objeto. Una onda convergente será reflejada con los cambios de fase que hemos descrito. En el punto focal, la onda incidente no será ni convergente ni divergente. Más allá del punto focal, la onda divergente tendrá un carácter diferente de reflexión. Las trazas osciloscópicas procedentes de la respuesta del transductor, en función del espaciado entre la lente y el objeto reflector, son reveladoras a este respecto. Abdullah Atalar, un investigador de nuestro grupo, ha mostrado que materiales con diferentes constantes elásticas, por ejemplo, producirán trazas osciloscópicas notablemente diferentes. La diferencia en el ángulo crítico para la reflexión total interna origina la diferencia en las porciones oscilantes de dichas trazas.

La respuesta del transductor para un substrato con una única capa superpuesta produce en el osciloscopio una traza particularmente interesante. Por ejemplo, hay un gran cambio en la forma de la traza correspondiente a una superficie de sílice pura cuando se deposita sobre ella una capa de aluminio de un micrometro. El cambio es todavía mayor si la capa depositada tiene dos micrometros de espesor [*véase la ilustración de esta página*]. En este ejemplo hay un espaciado lente-objeto correspondiente a un pico en la respuesta para la capa de dos micrometros y, en cambio, se presenta un valle en la respuesta para una capa de un micrometro. Con este método son naturalmente fáciles de acusar pequeños cambios en el espesor o en las propiedades elásticas de una película delgada. Por ejemplo, Robert G. Wilson y Rolf D. Weglein, de los Laboratorios de Investigación Hughes, han encontrado que el proceso acústico da lugar a notables variaciones de contraste en imágenes de objetos con múltiples capas, tales como los circuitos microelectrónicos.

Como hemos mencionado, la longitud de onda con la que opera el mejor microscopio acústico es aproximadamente de un micrometro, que no difiere mucho de las longitudes de onda de la luz visible. Hasta cierto punto puede decirse que se trata de una mera coincidencia. Si la longitud de onda límite para el sistema sonoro hubiera sido de diez micrometros, el microscopio acústico hubiera llamado poco la atención excepto para aplicaciones especiales. Si la longitud de onda límite hubiera sido de una décima de micrometro, la microscopía acústica habría sido preponderante. La importancia de un microscopio, por encima de todos los demás factores, queda definida por la resolución límite. En el microscopio acústico, ¿qué es lo que determina dicho límite? Se trata de la absorción del sonido a lo largo del camino acústico seguido en el líquido. En alguna forma hay que emplear líquidos si el objeto debe moverse respecto a la lente. Aunque los líquidos hacen posible el uso de longitudes de onda cortas, absorben energía. Las moléculas de muchos líquidos muestran modos de vibración que son excitados por las ondas de alta frecuencia. La absorción para tales líquidos es tan intensa que su empleo está excluido.

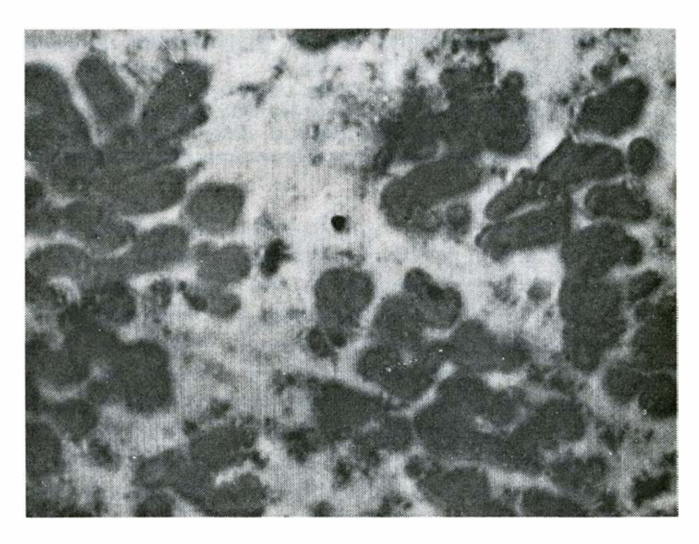
Con mucho, el líquido preferido ha sido el agua. Por fortuna, el agua es compatible con las muestras biológicas vivas. La velocidad de las ondas sonoras en el agua es de unos 1500 metros por segundo y la absorción es menor que en la mayoría de los otros líquidos, aunque aumenta con las frecuencias elevadas.





SUPERFICIE DE ACERO vista en una micrografía acústica (*derecha*). Es completamente diferente de su oponente óptica (*izquierda*). El mayor contraste entre granos cristalinos individuales en la imagen acústica debe ser atribuido a ligeras variaciones en las propiedades elásticas de los granos.





ALEACION de cobalto y titanio. Aparece en estas micrografías contrapuestas, una hecha con el microscopio óptico (*izquierda*) y la otra con el microscopio acústico (*derecha*). Las formas manchadas en la imagen acústica corresponden a regiones en las que la aleación tiene composición diferente.

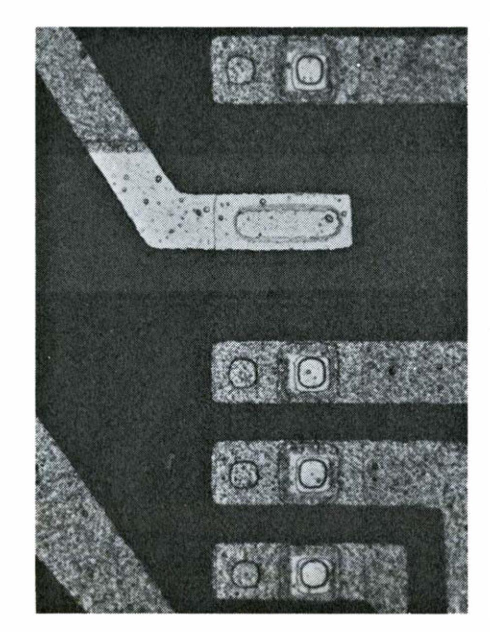
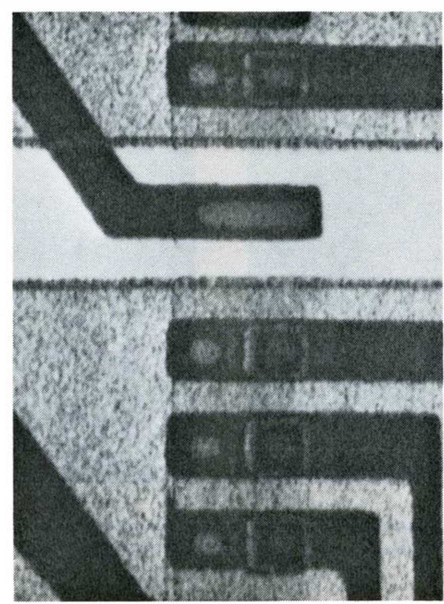
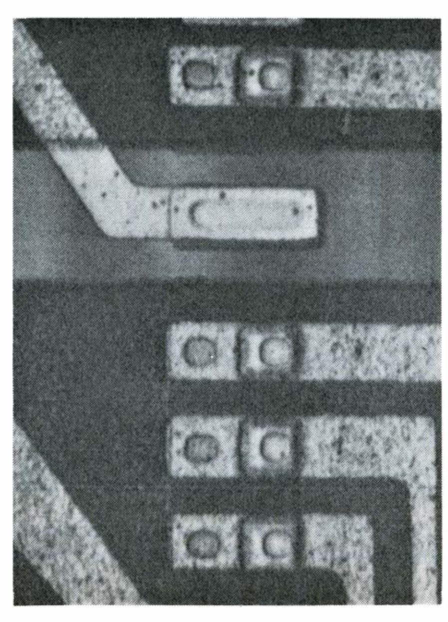
El argón es otro líquido interesante para la microscopía acústica. Sus propiedades acústicas son semejantes a las del agua excepto en cuanto a la velocidad del sonido, pues el argón tiene una velocidad que es sólo el 60 por ciento de la del agua. El argón tiene otro inconveniente: sólo existe en estado líquido en un estrecho margen de temperaturas en torno a los 85 grados Kelvin (menos 188 grados Celsius). Por tanto, es un líquido criogénico y hay que diseñar un equipo especial para los instrumentos que operan con tales líquidos. Así, se ha hecho para el microscopio acústico y Daniel Rugar, otro investigador de nuestro grupo, ha obtenido imágenes de elevada calidad con tal sistema.

El argón líquido es también importante por constituir un escalón intermedio en el paso de agua a helio líquido. La absorción del sonido en la fase líquida del isótopo helio 4 es tan pequeña que puede considerarse despreciable. La velocidad del sonido en tal medio es la octava parte de la del agua. En nuestro grupo, Joseph E. Heiserman está diseñando un microscopio acústico que emplee helio líquido como medio transmisor del sonido. Tiene la esperanza de conseguir un poder de resolución superior al del microscopio óptico.

Toda esta información de base constituye el prólogo de las imágenes conseguidas. En el pasado, a los visitantes a nuestro laboratorio les eran mostradas, en primer lugar, las imágenes y entonces se les preguntaba si deseaban ver el instrumento. Esta manera de proceder nos ahorraba mucho tiempo. En aquella época las micrografías acústicas poseían una calidad inferior a la de las micrografías ópticas y eran examinadas como una mera curiosidad. Los visitantes rara vez pedían darse una vuelta por el laboratorio y ver el instrumento en funcionamiento. Esto ya no ocurre. La mejora en calidad de las imágenes ha producido un intenso interés por el instrumento y nos vemos obligados a emplear cierto tiempo en explicar los detalles sobre su modo de operar.

Pero la esencia de cualquier programa de microscopía estriba en conseguir las imágenes. Creemos que las acústicas que acompañan el presente artículo pueden competir con sus oponentes ópticas. Las imágenes son de tres clases: de materiales, de circuitos de microelectrónica y de células vivas. El elemento común es el tamaño: todos los objetos tienen detalles cuyas dimensiones se miden en micrometros.

A simple vista, metales como el acero parecen ser uniformes, pero en realidad tienen estructuras muy complejas, que consisten en granos cristalinos individuales cementados conjuntamente a lo

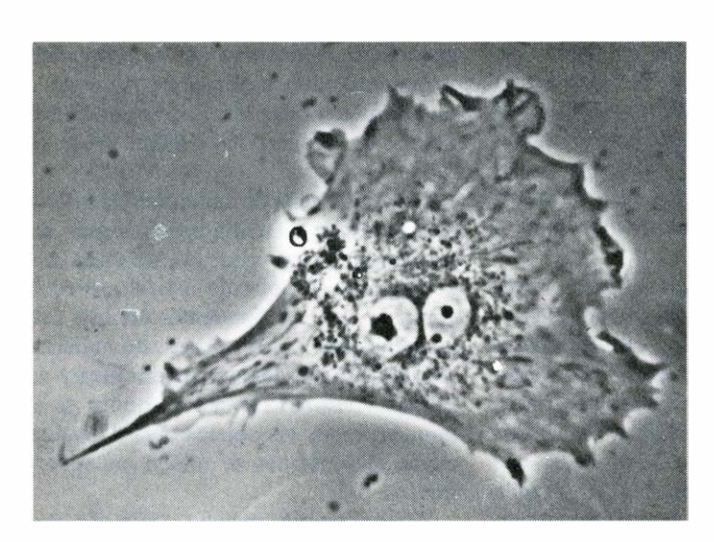
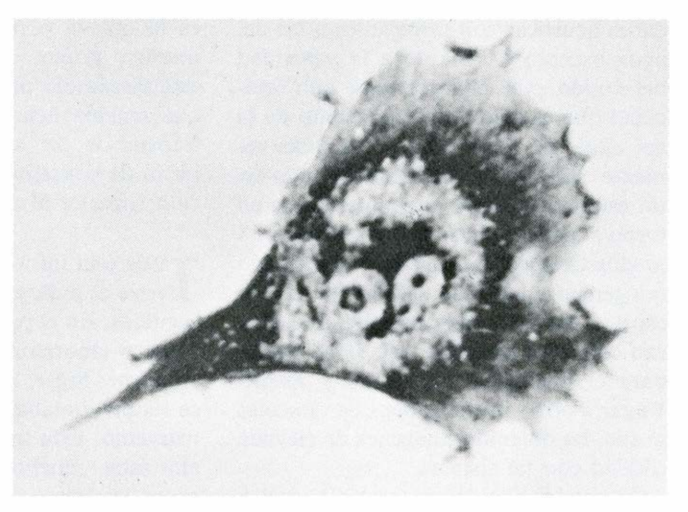
ELEMENTOS DE CIRCUITOS MICROELECTRONICOS construidos sobre la superficie de una pastilla de sílice. Se han reproducido tres veces: una en forma de una micrografía óptica (*izquierda*) y dos en forma de micrografías acústicas obtenidas con el objeto en diferentes posiciones focales (*centro y derecha*). El dispositivo consta de varios transistores interconectados. Las extensiones semejantes a dedos que conectan los transistores son tiras de películas de aluminio depositadas a partir del vapor. La banda lisa que cruza la parte superior de la imagen es el sustrato de sílice pura. La extensa textura en guijarros es ocasionada por capas superpuestas de sílice policristalina. La inversión de contraste en las imágenes acústicas ayuda a identificar la estructura detallada sobre la pastilla. El dispositivo fue proporcionado por el doctor Roderick D. Davies, de la Universidad de Stanford.

largo de sus bordes. La micrografía acústica de una superficie de acero parece completamente distinta de su oponente óptica [*véase la ilustración superior de la página precedente*]. El mayor contraste en la imagen acústica da una mejor idea de la orientación relativa de los cristales en la superficie. En nuestro caso, la superficie ha sido pulida y atacada de manera que los límites de grano parecen limpios en la micrografía óptica, pero este paso no es necesario con el sistema acústico.

Las aleaciones presentan otra clase de falta de homogeneidad. En una aleación de cobalto y titanio, por ejemplo, la mezcla puede condensarse en varias fases, cada una con una proporción entre cobalto y titanio distinta. Las fases difieren en su elasticidad, divergencia que constituye la causa del contraste en la imagen acústica [*véase la ilustración inferior de la página precedente*]. Las propiedades de estas aleaciones dependen grandemente del tamaño y la distribución de sus fases.

La circuitería microelectrónica constituye otra buena fuente de objetos para la micrografía acústica. Los detalles estructurales de tales dispositivos a menudo aparecen con mejor contraste que en las micrografías ópticas, que se ven más uniformes [*véase la ilustración superior de esta misma página*]. Además, con el instrumento acústico se pueden obtener varias imágenes diferentes del mismo objeto, cada una correspondiente a un espaciado distinto entre el objeto y la lente. Cada una de dichas imágenes proporciona un nuevo nivel de detalle. En nuestra investigación dicha información "en profundidad" ha motivado un interés más destacado que cualquier otra característica.

En el mundo biológico existen mu-

CELULA AISLADA. Se trata de un fibroblasto obtenido del tejido conjuntivo de un embrión de pollo. Se ha representado mediante su imagen óptica (*izquierda*) y la correspondiente imagen acústica (*derecha*). La célula se hizo crecer sobre un portaobjetos de vidrio con colágeno y fue fijada con metanol por Randal N. Johnston en Stanford. Ambas micrografías revelan la red de fibras internas que afectan a toda la célula. El contraste en la imagen óptica se obtuvo con una técnica de tinción selectiva. El contraste en la imagen acústica procede de variaciones en las propiedades elásticas de las estructuras en la célula. El que la técnica acústica no requiera tinción sugiere que podemos llegar a visualizar tales estructuras en células vivas.

chas estructuras celulares en la escala de los micrometros. De una manera rutinaria se visualizan con el microscopio óptico, pero no sin dificultad. A menudo se presenta una falta de contraste visible. El problema reside en que las estructuras de las células son casi completamente transparentes: sólo absorben una minúscula parte de luz. Se han realizado extraordinarios esfuerzos para compensar tal deficiencia y aumentar el contraste. El método usual es el del teñido selectivo de las células. Constituye un arte altamente desarrollado y añade color a las imágenes, pero requiere tiempo y no puede ser aplicado a las células vivas. Las técnicas ópticas asociadas con la microscopía de contraste de fase y la microscopía de luz polarizada permiten introducir cambios sutiles para aumentar el contraste, pero todavía queda campo para introducir nuevas mejoras.

Al emprender esta parte de la investigación, nuestra ignorancia entonces en las cuestiones biológicas quedará ilustrada con una anécdota. Cuando por primera vez preguntamos la manera de obtener glóbulos blancos de la sangre, en un laboratorio biológico de Stanford se nos pidió que dijéramos cuántos glóbulos deseábamos. Ya que queríamos obtener unas pocas imágenes, contestamos que necesitamos diez glóbulos. Esta respuesta sólo podía proceder de un profano en biología. Se nos explicó que el número mínimo de glóbulos que podía aislarse era el de un millón y todos cabían en un simple portaobjetos. En muchos casos sería deseable estudiar el conjunto de estos glóbulos, pero esto no constituye una cuestión fácil. El coste de la película fotográfica necesaria para registrar las imágenes de todos estos glóbulos sería impresionante. Una de tales exposiciones cuesta unas tres pesetas y en los laboratorios de biología es corriente registrar unas veinte imágenes de cada glóbulo. Por tanto, el coste de la película sería aproximadamente de unas 65 pesetas por glóbulo. Incluso si se limitase el estudio a sólo el uno por ciento de los glóbulos, el coste (650.000 pesetas) resultaría prohibitivo. Con ayuda de circuitos especiales conectados a nuestro instrumento acústico puede reducirse dicho coste. Pronto estaremos capacitados para registrar tales imágenes directamente en cinta magnética. El tiempo de reproducción en un registrador magnetofónico (a una cadencia de treinta imágenes por segundo) para 10.000 glóbulos sería de una hora y el coste de la cinta sería inferior a unas 1200 pesetas.

Hasta ahora hemos realizado micrografías de tres clases principales de entidades biológicas: linfocitos (glóbulos blancos de la sangre), fibroblastos (células de soporte de los tejidos) y cromosomas. El hecho de haber obtenido imágenes tan buenas de distintas estructuras celulares sin necesidad de emplear tinciones indica que podremos observar algunas de tales estructuras en el interior de las células vivas.

Las imágenes que hemos elegido para ilustrar el presente artículo son ideales en muchos aspectos para la microscopía acústica y demuestran lo que todavía puede conseguirse. Son muy diferentes de las imágenes que era obvio elegir cuando emprendimos este tipo de estudio. Cuando echamos una mirada hacia atrás comprobamos que los ejemplos primitivos, tales como la unión entre materiales opacos, fueron simples extensiones del trabajo que se venía haciendo acerca de la formación acústica de imágenes a frecuencias ultrasónicas más bajas. Los ejemplos que mostramos no estaban incluidos en nuestra lista original. Han evolucionado de una manera tan natural como inesperada.

En una ocasión estábamos seguros de que la reflexión de las ondas sonoras en una superficie de zafiro puro sólo debía venir determinada por el contorno del haz acústico. Se presentaría un máximo al colocar el reflector en el punto focal que iría disminuyendo de acuerdo con el perfil del haz a medida que el reflector fuera separado del plano correspondiente. Nada nos hacía suponer que las propiedades elásticas de la superficie reflectora quedaban codificadas en las ondas divergentes. En otra ocasión, trabajamos intensamente para perfeccionar un microscopio acústico del tipo de transmisión. Nunca sospechamos que las imágenes obtenidas por reflexión, particularmente de muestras biológicas blandas, pudieran ser tan vigorosas.

Nuestro futuro programa de investigación se plantea en forma de una serie de preguntas. ¿Hasta dónde podemos llegar con los líquidos criogénicos? ¿Se hallará un líquido poco absorbente, no criogénico, que permita alcanzar longitudes de onda más cortas y resoluciones más elevadas a la temperatura ambiente? ¿Se atravesará la barrera sólido-líquido y podremos enfocar el microscopio hacia el interior de materiales sólidos? ¿Seremos capaces de conseguir imágenes de líneas de esfuerzos en tales materiales? ¿Se desarrollarán células vivas en el microscopio acústico y darán información significativa para los biólogos? Es como si acabáramos de comenzar.

Ciencia y sociedad

Lluvia ácida

La acidez en una solución tal como la lluvia es sinónimo de la presencia de iones hidrógeno (H^+). Una medida corriente de la acidez es el *p*H, que se define como el logaritmo cambiado de signo de la concentración de iones hidrógeno. La escala de *p*H varía desde 0 a 14, con el valor 7 representando una solución neutra mientras que los inferiores a 7 indican acidez creciente y los superiores a 7 alcalinidad creciente. Debe tenerse en cuenta que la escala de *p*H es logarítmica, de modo que las soluciones de *p*H 6, 5 y 4 contienen respectivamente uno, 10 y 100 C microequivalentes de acidez (H^+) por litro (abreviado meq/l).

¿Qué determina la química de la lluvia y la nieve? El agua proviene de la evaporación y la transpiración (vapor de agua perdido por las plantas) y es esencialmente destilada, o pura. Una vez el vapor alcanza la atmósfera, sin embargo, se condensa sobre partículas sólidas y pronto alcanza el equilibrio con los gases atmosféricos. Uno de los gases es el dióxido de carbono (CO_2), y a medida que el dióxido de carbono se disuelve en el agua, se forma ácido carbónico (H_2CO_3). Como el ácido carbónico es débil, se disocia sólo ligeramente en agua destilada, dando iones hidrógeno e iones bicarbonato (HCO_3^-). A presiones y concentraciones normales de dióxido de carbono en la atmósfera, el *p*H de la lluvia y nieve sería de 5,6.

Otras substancias que llegan a la atmósfera tienden a alterar el *p*H en un sentido o en otro. Por ejemplo, pequeñas cantidades de polvo y residuos son barridas por el viento y pasan del suelo a la atmósfera. Las partículas de suelo son en general ligeramente básicas o alcalinas, y al disolverse en agua destilada liberan en la solución cationes básicos (iones positivos), tales como calcio, magnesio, potasio y sodio (Ca^{++}, Mg^{++}, K^+ y Na^+), siendo generalmente el bicarbonato el correspondiente anión o ion negativo. El gas amoníaco en la atmósfera se origina en gran parte como resultado de la degradación de materia orgánica. El ion amonio (NH^{2+}) en la lluvia y nieve tiende a aumentar el *p*H. En las regiones costeras, la espuma marina puede contribuir en forma importante a la química de la precipitación. Aquí los iones principales que se añaden son los más abundantes en agua de mar, esto es, los cationes Na^+ y Mg^{++} (sodio y magnesio), el anión Cl^- (cloro) y en menor grado cationes de calcio y de potasio (Ca^{++}, K^+) y aniones de sulfato (SO_4^{--}).

Los gases tales como el dióxido de azufre (SO_2) y el sulfuro de hidrógeno (H_2S), que provienen de los volcanes y otras fuentes, pueden también alterar localmente la química de la precipitación. El dióxido de azufre y el sulfuro de hidrógeno se oxidan e hidrolizan en la atmósfera pasando a ácido sulfúrico. Los óxidos de nitrógeno se convierten de manera análoga en ácido nítrico. Si esos ácidos están presentes en cantidades apreciables, pueden acidificar la precipitación por debajo del *p*H 5,6.

La química de la precipitación natural depende, por tanto, de las cantidades relativas de esas diversas substancias en la atmósfera. La precipitación en regiones costeras se puede generalmente caracterizar como agua del mar muy diluida si los vientos soplan hacia la costa. En regiones de vientos fuertes y vegetación escasa, tales como los desiertos, la precipitación puede contener cantidades relativamente grandes de cationes básicos. La precipitación en regiones con suelos calcáreos contiene frecuentemente calcio y bicarbonato, presumiblemente como resultado de la incorporación de polvo, y el *p*H de la lluvia y nieve es generalmente bastante superior a 6.

El *p*H de la precipitación que caía antes de la Revolución Industrial, y que se ha conservado en glaciales y barreras de hielo continentales, era superior a 5. Recientemente, investigadores suizos determinaron en Groenlandia el *p*H de hielo que se había originado en forma de nieve hace 180 años; encontraron *p*H que variaba entre 6 y 7,6.

Las actividades humanas han alterado apreciablemente este cuadro en una escala continental y quizá global. Se emiten a la atmósfera grandes cantidades de óxidos de azufre y nitrógeno al quemar combustibles fósiles y en la tostación de sulfuros, particularmente en la zona templada septentrional, fuertemente industrializada y urbanizada. Esos óxidos se convierten en ácidos fuertes (sulfúrico y nítrico) en la atmósfera y caen al suelo en forma de lluvia y nieve. Los ácidos fuertes se disocian completamente en soluciones acuosas diluidas y hacen bajar el *p*H a menos de 5,6. La definición de precipitación ácida es lluvia y nieve con un *p*H por debajo de ese valor.

Una ironía en esta cuestión es que la tendencia a construir chimeneas cada vez más altas para mitigar los problemas de contaminación local ha convertido los problemas locales en regionales. Durante el pasado decenio, las emisiones anuales de azufre procedentes de Sudbury, Ontario, han igualado aproximadamente la cantidad que se cree han emitido todos los volcanes del mundo.

Está fuera de duda que los contaminantes son transportados a grandes distancias por la atmósfera. Estudios realizados sobre la deposición de azufre, plomo y otros materiales en el casquete de hielo de Groenlandia y en otros lugares remotos ponen de manifiesto tales mecanismos de difusión. Lo que es objeto de discusiones, a veces acaloradas, particularmente cuando se alega que la contaminación atraviesa fronteras, es la proporción de contaminantes que puede atribuirse a fuentes remotas.

Tomando como base el año, la lluvia y la nieve sobre grandes regiones del mundo son ahora entre cinco y treinta veces más ácidas que el valor más bajo (*p*H 5,6) que debe esperarse para atmósferas no contaminadas. La lluvia procedente de borrascas individuales puede ser de varios cientos a varios miles de veces más ácida de lo que debería.

El descubrimiento de que la precipitación en grandes extensiones de Europa se había vuelto decididamente ácida siguió al establecimiento en el decenio de 1950 de la Red Europea de Química Atmosférica (también conocida como Red IMI, por el Instituto Meteorológico Internacional de Estocolmo). En ciertos períodos, la red abarcaba hasta 175 estaciones en Europa septentrional y occidental, de la cual se analizaban muestras mensuales de la precipitación total (método que está al alcance de todo el mundo) en lo referente a las concentraciones de 11 iones importantes.

La Organización de Cooperación y Desarrollo Económico (OCDE) finalizó recientemente un estudio de tres años que da un cuadro más detallado de la actual acidez de la precipitación en Europa. El mapa resultante muestra que la precipitación ácida ha continuado extendiéndose y hoy abarca casi todo el Noroeste de Europa.

En algunas estaciones, el aumento de la concentración anual de iones hidrógeno en la precipitación desde 1955-56 se ha multiplicado aproximadamente por diez. El aumento de acidez es comparable con los aumentos de concentración de los sulfatos y nitratos a lo largo de los pasados veinte años. En el Norte de Europa, el exceso de concentración de sulfatos (exceso sobre la contribución

de la sal marina) aumentó marcadamente de 1955 hasta aproximadamente 1963, y desde entonces ha permanecido virtualmente constante; en conjunto, el aumento de 1955 a 1975 ha sido de aproximadamente dos por ciento por año. La concentración de nitrato en la precipitación ha aumentado en forma muy constante desde 1955 al ritmo de aproximadamente un cuatro por ciento anual. Ambos aumentos corren paralelos al aumento de las emisiones de óxidos de azufre y nitrógeno producidos al quemar combustibles fósiles.

El aspecto general de la deposición de ácido en las lluvias y nieves europeas refleja el transporte a gran distancia de los contaminantes atmosféricos desde las fuentes de emisión. El *p*H de la precipitación es generalmente mínimo cerca de las fuentes y aumenta con la distancia. La dirección normal de las corrientes atmosféricas de suroeste a nordeste da lugar a una distribución asimétrica, con la precipitación ácida alcanzando muy hacia el norte en Escandinavia. El gradiente de acidez se acentúa en áreas montañosas que están generalmente viento abajo de las fuentes urbanas e industriales. Las montañas favorecen la precipitación, que continuamente elimina los contaminantes de las masas de aire que pasan sobre ellas. Para la época en que las masas de aire han pasado sobre las áreas montañosas con una elevación de unos 1000 metros, el *p*H de la precipitación ha variado de un promedio de 4,2 en la costa a 4,6 o 4,7, lo que representa una disminución en la concentración de iones hidrógeno desde unos 60 microequivalentes por litro a unos 20. Como las montañas favorecen la precipitación, sus laderas de barlovento reciben grandes cantidades de ácido. En las laderas montañosas del extremo Sur de Noruega, Alemania meridional y Suiza se observaron valores que variaban entre 88 y 120 miliequivalentes por metro cuadrado y año.

La deposición seca de gas dióxido de azufre y sulfato en forma de partículas representa también la deposición de una cantidad equivalente de iones hidrógeno. Los datos del estudio de la OCDE muestran valores máximos de deposición seca de 10 gramos de azufre por metro cuadrado y año en Europa central, lo que es cuatro veces el ritmo de deposición por lluvia y nieve, con ritmos que disminuyen rápidamente al aumentar la distancia a las fuentes de emisión hasta valores de aproximadamente 0,1 gramos en el Norte de Escandinavia. La deposición húmeda de sulfatos en la lluvia y nieve empieza a tres gramos de azufre por metro cuadrado y año cerca de las fuentes más importantes y disminuye menos rápidamente hasta 0,5 gramos en el Norte de Escandinavia. La deposición seca de azufre es, por tanto, el modo dominante cerca de las fuentes de emisión, mientras que la deposición húmeda domina en áreas remotas del Norte de Europa.

¿Qué efecto ha producido el aumento de acidez sobre el medio ambiente? No hay todavía una respuesta completa para la pregunta, pero ciertos ecosistemas de agua dulce parecen ser particularmente sensibles al ingreso de ácido. Tales ecosistemas se encuentran generalmente en regiones donde el lecho rocoso subyacente está formado por rocas de tipo altamente silíceo, tales como el granito, algunos gneiss, la cuarcita y la arenisca de cuarzo. Estos tipos de rocas son fuertemente resistentes a la disolución por efecto de la intemperie, con el resultado de que las aguas superficiales en tales regiones contienen concentraciones muy bajas de los iones que se producen debido a la erosión por la intemperie. Característicamente, tales aguas blandas tienen muy baja capacidad para neutralizar la adición de ácidos o álcalis. De aquí que cuando la precipitación ácida cae sobre tal región, los ácidos no quedan completamente neutralizados en la cuenca terrestre y los arroyos y lagos se acidifican.

La geología del fondo rocoso, por consiguiente, suministra una grosera indicación de la localización de tales áreas sensibles. Los mapas del lecho rocoso, sin embargo, no toman en consideración la mineralogía de los depósitos sedimentarios situados encima, tales como material de arrastre de glaciares, materiales glaciofluviales y arena marina, que pueden haber sido transportadas desde muy lejos de su fuente y tener una mineralogía enteramente diferente de la del lecho rocoso. Pueden darse variaciones en la química de las aguas superficiales como resultado de diferencias en esos dos substratos geológicos, en el tipo de suelo y en la vegetación.

Las regiones de lechos rocosos altamente silíceos están muy extendidas en el Escudo fennoscándico precámbrico en Escandinavia, el Escudo canadiense, las Rocosas, Nueva Inglaterra, los Adirondacks, los Apalaches y regiones menores en otros lugares. Las aguas dulces acidificadas hasta *p*H inferiores a 5 como resultado de precipitación ácida se encuentran en Noruega meridional, Sur de Suecia, Gran Bretaña, Holanda, Alemania Oriental, Nueva Escocia, Ontario, parte Norte de Nueva Inglaterra, los Adirondacks y Florida. Se ha producido un efecto claro sobre las poblaciones de

peces, que han sido virtualmente barridas o fuertemente reducidas en los lagos acidificados de algunas de esas regiones. Además, se han observado otros varios cambios ecológicos como resultado del incremento de ácido en tales lagos. La velocidad de descomposición de materia orgánica se ha hecho más lenta, presumiblemente como consecuencia de cambios en las bacterias y hongos predominantes. Las bacterias, que son eficaces en la descomposición, generalmente se comportan peor en un ambiente ácido. Se producen cambios en todos los niveles de la cadena de alimentación: el fitoplancton, el zooplancton y los peces, todos los cuales disminuyen en número de especies. Incluso el comportamiento de los organismos puede cambiar como consecuencia de cambios en el medio ambiente.

La distribución de lagos ácidos concuerda con la distribución de precipitación ácida. Ulterior prueba del impacto de la precipitación ácida es que las aguas dulces ácidas no se encuentran normalmente en regiones que serían sensibles al ácido (por ser pobres en nutrientes y poco aptas para neutralizarlo), sino que quedan fuera de las regiones de precipitación fuertemente ácida. Estas observaciones conducen a la conclusión de que tanto la cantidad de ácido en la precipitación como las características geológicas de la zona son importantes para determinar la sensibilidad de las aguas dulces a la precipitación ácida.

Parece improbable que el consumo de combustibles fósiles en los países industrializados disminuya en los próximos 30 años o así. Es más probable que aumente a causa de la creciente tendencia a reemplazar el petróleo por carbón. Debe hacerse frente al problema de los efectos ambientales en gran escala.

La instalación de desulfurizadores por vía húmeda (scrubbers) es cara, pero los costes ambientales de emisiones inadecuadamente controladas son también elevados. Muchos costes están ocultos, particularmente los asociados con los efectos sinérgicos de las tensiones que el ácido impone a los organismos (por ejemplo, aumentar su susceptibilidad a patógenos y depredadores). Las soluciones sencillas, tales como quemar petróleos y carbones con bajo contenido de azufre, se han aplicado ya en el control de emisiones en algunas regiones. Sin embargo, en algunas zonas el *p*H de la lluvia y nieve ha sobrepasado ya la parte "fácil" de la escala (de 7 a 4). La reducción de la acidez será ahora más difícil y mucho más costosa de llevar a cabo. En último término, la solución del problema de la precipitación ácida estriba en el decrecimiento de las emisiones de óxidos de azufre y nitrógeno a la atmósfera. Este objetivo se puede lograr mediante una combinación de mejora de la tecnología de control y de conservación de combustibles.

Nobel de física

El premio Nobel de física correspondiente al año 1979 se ha concedido conjuntamente a tres físicos teóricos: dos norteamericanos, Sheldon L. Glashow y Steven Weinberg, y un pakistaní, Abdus Salam. Los primeros trabajan en la Universidad de Harvard, el tercero simultanea su puesto en el Imperial College de Londres con el de director del Centro Internacional de Física Teórica de Trieste. El motivo de concesión del premio ha sido común para los tres: sus contribuciones a las teorías unificadas de las interacciones fundamentales [véase, para una mayor profundización el artículo del autor de esta nota: "Teorías unificadas de las interacciones fundamentales", por Francisco J. Ynduráin, en *Investigación y Ciencia,* marzo, 1978].

Las interacciones débiles son conocidas experimentalmente desde hace bastante tiempo (Becquerel, 1896); su primera descripción téorica fue propuesta por Fermi en 1932. La hipótesis de Fermi consistía en suponer una interacción entre cuartetos de fermiones en la que un par se aniquilaba a la vez que se creaba un nuevo par (interacción de contacto). El problema con esta interacción era que, a pesar de su valor fenomenológico, daba resultados inconsistentes ("divergencias") en cuanto se intentaba aplicar en órdenes superiores al más bajo en teoría de perturbaciones. Lee y Yang intentaron reducir el problema de las divergencias, suponiendo que los pares de fermiones que intervenían en las reacciones débiles se agrupaban en corrientes similares a las electromagnéticas, y que la interacción no era de contacto sino transmitida por unos hipotéticos campos, los bosones *W*, similares al fotón como transmisor de las interacciones electromagnéticas (excepto que se las suponían muy masivas). Esta analogía llevó a Schwirger y, sobre todo, a Glashow, a proponer un modelo en el que las interacciones débiles y las electromagnéticas aparecían unificadas.

El modelo de Glashow no fue tomado muy en serio al demostrar Salam y otros que contenía problemas de consistencia iguales (aunque algo menos severos) que la interacción original de Fermi. Posteriormente, Weinberg (1967) y el propio Salam (1968) y Ward mejoraron el modelo de Glashow introduciendo las masas de los *W*, y la unificación de interacciones débiles con electromagnéticas, a través del llamado "mecanismo de ruptura espontánea de simetría". Tampoco estos modelos mejorados fueron muy populares hasta que, en 1971, G. 't Hooft demostró que estaban libres de inconsistencias y podían, por tanto, constituir una teoría satisfactoria unificada para las interacciones débiles y electromagnéticas.

En un principio, el modelo fue formulado sólo para partículas ligeras (leptones); su extensión a partículas nucleares (quarks) presentaba problemas. Una forma de salvar dichos problemas era la introducción de un nuevo número cuántico, el famoso encanto, debida a Bjorken y Glashow (1964), a Glashow, Iliopoulos y Maiani (1970) y a Weinberg (1972). Con esto la teoría contenía dos predicciones experimentalmente comprobables: la existencia de corrientes débiles neutras, y la del encanto. La primera fue verificada experimentalmente en el CERN (Ginebra), en primer lugar, en 1972/1973 con el descubrimiento del famoso suceso de colisión elástica $\gamma_\mu + e \longrightarrow \gamma_\mu \pm e$, y luego confirmada en numerosos experimentos. La segunda se comprobó de manera espectacular al descubrir un año después Richter y Tirg las llamadas partículas ψ/J (por lo que recibieron el premio Nobel en 1976) cuya interpretación teórica en términos de encanto fue dada por Appelquist y Politzer y por De Rújula y Glashow en 1975. En el momento actual, el modelo que, con cierta injusticia, se ha popularizado con el nombre de modelo de Weinberg-Salam, es una de las más afortunadas teorías existentes, en el sentido de capacidad predictiva.

El modelo se ha extendido también hasta la unificación de las llamadas interacciones fuertes por Pati y Salam y por Georgi y Glashow. Curiosamente, la única interacción que queda fuera del esquema, en el sentido de no existir modelos que la incorporen y con la doble virtud de ser teóricamente atractivos y claramente compatibles con los resultados experimentales, es la interacción gravitatoria: precisamente la que originalmente Einstein intentó tomar como base para una unificación.

La concesión de dos premios Nobel a físicos de partículas (un grupo de experimentadores y otro de teóricos) en el corto intervalo de tres años refleja los espectaculares hallazgos experimentales y desarrollos teóricos de los años 70 que arrancaron a este área de la ciencia del punto muerto en que se encontraba, abriendo a nuestra comprensión fenómenos fundamentales. (F.J.Y.)

Legionelosis

La misteriosa enfermedad epidémica de los legionarios, aparecida en 1976, se atribuyó a una bacteria hasta entonces desconocida. Se ha comprobado que la legionelosis, infección causada por dicha bacteria, se presenta con frecuencia

David W. Fraser y Joseph E. McDade

La LVIII convención anual del Departamento de Pennsylvania de la Legión Americana se celebró en el hotel Bellevue-Stratford, del 21 al 24 de julio de 1976. En muchos aspectos, fue una convención más, a la que asistieron 4400 delegados, familiares y otros participantes. Hubo parada militar, conferencias y los actos, protocolarios e informales, que son habituales en esos acontecimientos. Nadie pensaba entonces que la convención pasaría a los anales de la medicina. Entre el 22 de julio y el 3 de agosto, 149 participantes de la convención comenzaron a mostrar síntomas de una extraña enfermedad, caracterizada, en todos los casos, por accesos de fiebre, tos y neumonía. Esta coincidencia suponía un brote espectacular e insólito de neumonía sin causa aparente. Puesto que muchos de los reunidos habían vuelto ya a sus hogares antes de sentirse enfermos, el Departamento de Sanidad de Pennsylvania no dio cuenta del brote entre los asistentes hasta el 2 de agosto.

Ese día marcó el inicio de una de las investigaciones más amplias y complejas desarrolladas jamás en el esclarecimiento de los orígenes de una epidemia. La enfermedad de los legionarios, como la bautizó la prensa de inmediato, planteaba un formidable reto a epidemiólogos e investigadores en general. Se tardó varios meses en descubrir el agente desencadenante, y hubo de transcurrir aún bastante tiempo hasta lograr la confirmación de que dicho agente era una bacteria hasta entonces desconocida. Por de pronto, la enfermedad de los legionarios resultó ser menos rara e infrecuente de lo imaginado, responsable asimismo de otros brotes y de casos esporádicos en muchas partes del mundo, y ello antes y después del famoso episodio de Filadelfia.

En la investigación de una epidemia, la fase inicial consiste en determinar las características de la enfermedad, quién la ha contraído, dónde y cuándo. A continuación hay que establecer las circunstancias comunes y exclusivas de los aquejados de ella: dónde se encontraban y en qué difirió su comportamiento del resto del grupo que no la contrajo. El conocimiento de esos pormenores puede revelarnos el modo de propagación del agente causal y, por tanto, sugerir la identidad del mismo, así como su origen. Para reunir lo antes posible esa información sobre los participantes en la convención de Filadelfia, esparcidos por todo el estado de Pennsylvania, colaboraron con el personal sanitario municipal y estatal 23 funcionarios del Centro de Control de las Enfermedades (CDC, del inglés "Center for Disease Control"), del Servicio de Salud Pública de los Estados Unidos. Muy pronto advertimos que la enfermedad no se circunscribía a los legionarios. Se descubrieron otros 72 casos en individuos que no habían tenido relación directa con aquella convención militar. Pero sí mostraban algo en común con los legionarios: por una u otra razón habían estado en el hotel Bellevue-Stratford, o en sus proximidades. El número total de pacientes se elevó a 221, de los cuales 34 murieron de neumonía o a causa de sus complicaciones posteriores.

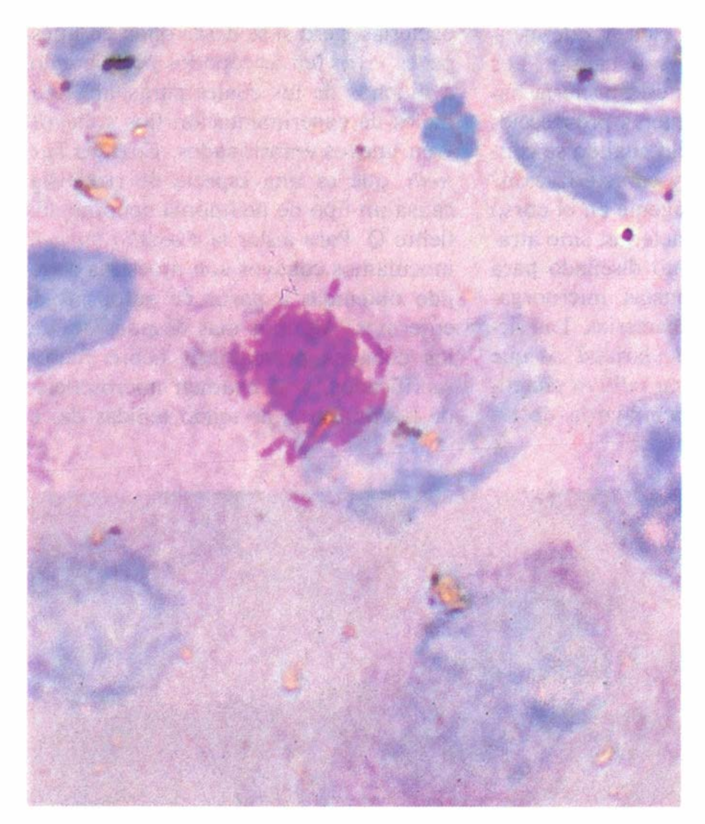
Epidemiología

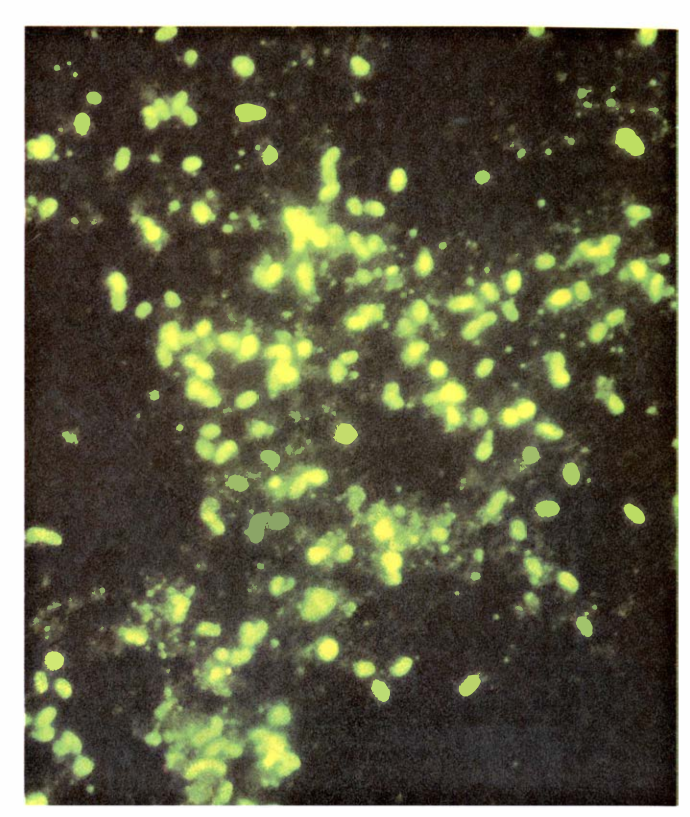
Los asistentes a la concentración militar de edad avanzada se vieron afectados en una proporción mucho mayor que los jóvenes, y los varones en una tasa tres veces superior que las mujeres. Los legionarios que se habían alojado en el Bellevue-Stratford mostraron una incidencia mucho más alta que los que pernoctaron fuera, y, entre estos últimos, quienes enfermaron habían permanecido más tiempo en el hotel que quienes no lo hicieron. Los únicos familiares que contrajeron la enfermedad también habían estado en Filadelfia con motivo de la convención, lo que ponía de manifiesto que la epidemia no se transmitía de persona a persona. La posibilidad lógica de que la infección se hubiera propagado a través de la comida o bebida quedó descartada. No se registraron diferencias, entre los afectados y los no afectados, por lo que se refería a restaurantes visitados, asistencia a banquetes, bebidas o hielo consumidos.

Búsqueda del agente

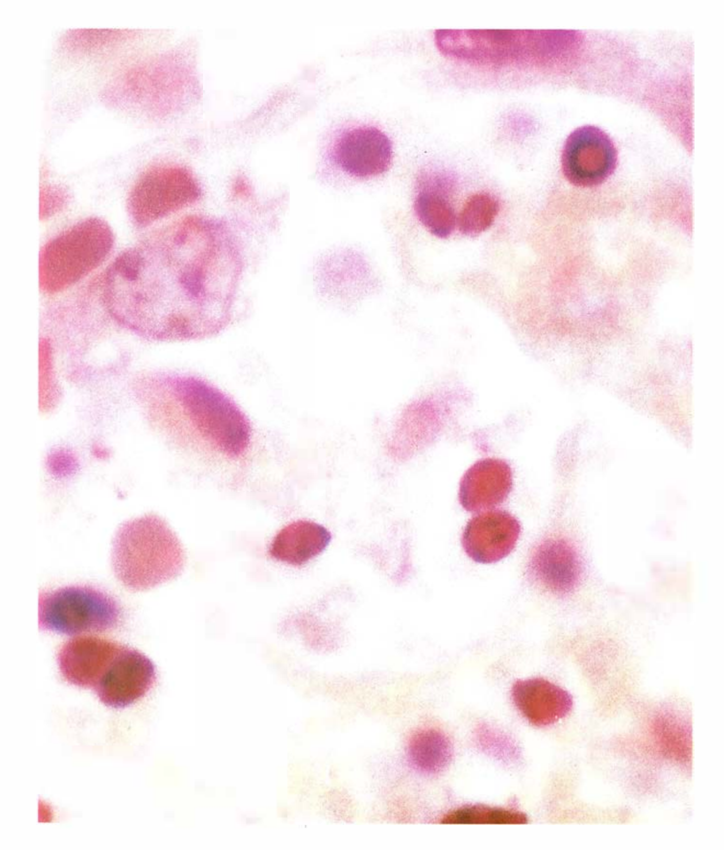
Ciertas observaciones sugirieron que la enfermedad podía haberse propagado por el aire. Los legionarios enfermos habían permanecido por más tiempo (alrededor del 60 por ciento más) en el vestíbulo de Bellevue-Stratford que los que no cayeron enfermos; asimismo, los legionarios afectados habían estado más tiempo en la acera del hotel que sus compañeros no atacados. Pero si el agente inductor de la enfermedad hubiera contaminado el aire del vestíbulo y hubiera salido a la acera por la puerta del vestíbulo, hubieran caído legionarios y clientes que se encontraran allí. Y como ese parecía ser el caso, lo más probable era que el sistema de transmisión fuera el aire. A pesar de todo, pasó todavía algún tiempo antes de saberse a ciencia cierta cuál había sido el agente causal que había desencadenado la galopante enfermedad de los legionarios.

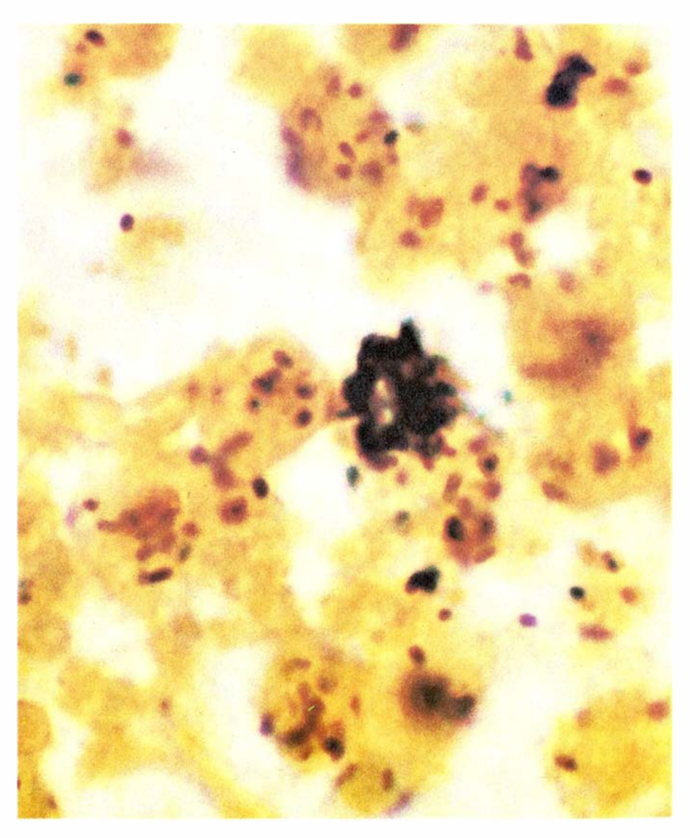
Los síntomas clínicos podían deberse a una amplia variedad de agentes, entre los que se contarían metales pesados, sustancias orgánicas tóxicas y organismos infecciosos. Durante las primeras semanas de investigación, las muestras obtenidas de los enfermos y los tejidos procedentes de autopsias fueron sometidos a intensivos análisis. La comparación entre muestras procedentes de enfermos y controles no parecía revelar una presencia significativa de niveles poco usuales de metales o sustancias tóxicas de origen orgánico en los enfermos que pudieran relacionarse con la epide-





AGENTE CAUSAL de la enfermedad de los legionarios. Se observó por primera vez en una preparación de bazo de un cobayo inoculado con tejido pulmonar de un enfermo que había fallecido víctima de dicha enfermedad. La microfotografía (*izquierda*) muestra un grupo de organismos alargados (bacilos) teñidos de rojo por el método de Giménez. Más tarde se demostró que estos organismos eran bacterias, cuando crecieron normalmente en un medio bacteriológico modificado. Por medio de ensayos de inmunofluorescencia indirecta se puso de manifiesto que habían causado la enfermedad de los legionarios. En un ensayo similar (directo) se tiñen suspensiones de tejido pulmonar de enfermos con un anticuerpo procedente de conejo y unido a fluoresceína, específico para el agente de la enfermedad de los legionarios. El anticuerpo se fija a la bacteria, que forma una coloración verde-amarillenta brillante al ser observada en un microscopio de luz ultravioleta (*derecha*). Ambas preparaciones se presentan ampliadas 2500 veces.





LAS TINCIONES CONVENCIONALES para la observación de bacterias en muestras patológicas (tejidos de un enfermo) no constituyen un buen método adecuado para la bacteria de la enfermedad de los legionarios. Una sección pulmonar teñida por la técnica de Brown y Brenn, de uso generalizado, no pone de manifiesto la presencia de organismos (*izquierda*). Francis W. Chandler, del Centro para el Control de las Enfermedades (CDC "Center for Disease Control"), descubrió que una versión modificada del método de impregnación con plata de Dieterle podía teñir la bacteria, como se muestra en la micrografía de la derecha, obtenida de una sección del mismo pulmón que en el caso anterior. Secciones ampliadas 2500 veces.

mia. No se encontró ningún agente microbiano conocido que pudiera explicar el brote. Aun así los análisis continuaron, prestándose especial atención a la posibilidad de que el responsable hubiera sido algún organismo desconocido, que de alguna forma hubiera logrado escapar a la detección. Esta suposición resultó ser correcta.

En enero de 1977, en colaboración con Charles C. Shepard y otros investigadores del CDC aislamos a partir de legionarios que habían muerto, víctimas de la enfermedad, una bacteria patógena hasta entonces desconocida, y que identificamos como el agente causal. Los motivos por los cuales esta bacteria había escapado a las primeras prospecciones quedaron claros sólo cuando se consiguió su aislamiento y caracterización.

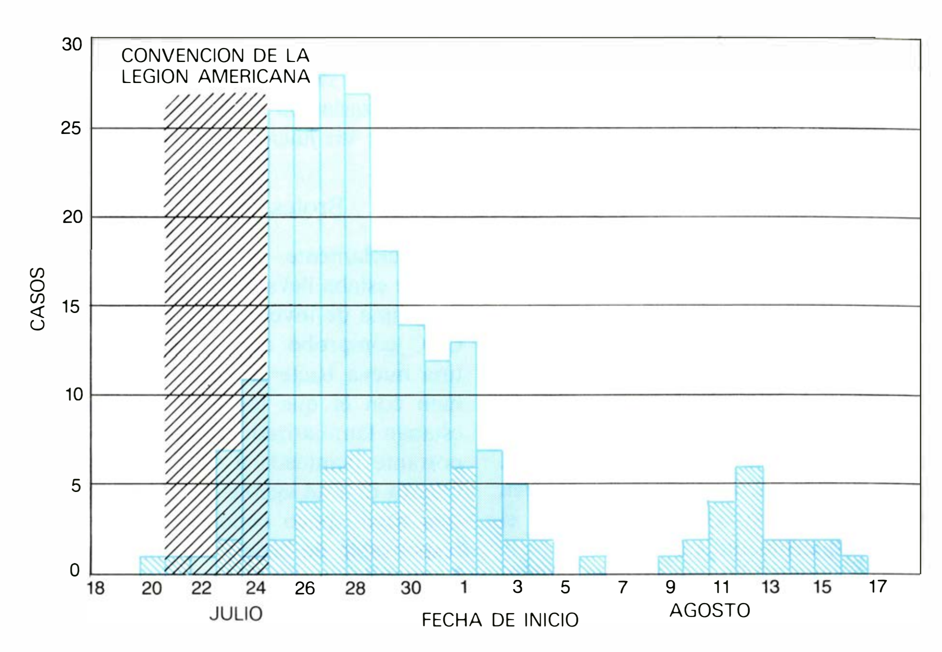
No descubrimos el agente en el curso de un aislamiento de bacterias, sino através de un procedimiento diseñado para el aislamiento de rickettsias, microorganismos similares a las bacterias. Las rickettsias tienen la particularidad de que no crecen en medios de cultivo sintéticos, como lo hacen la mayoría de las bacterias; pero sí se desarrollan en huéspedes vivos (en artrópodos por ejemplo, a muchos de los cuales parasitan), animales de experimentación (los cobayos) y en huevos embrionados. *Coxiella burnetti,* que es una especie de rickettsia, causa un tipo de neumonía denominada fiebre Q. Para aislar la *Coxiella burnetti* inoculamos cobayos con muestras de tejido obtenidas a partir de autopsias de enfermos. Algunos días después de que los cobayos presentaran fiebre fueron sacrificados, y el examen microscópico de las muestras de tejido teñidas dio el



***LEGIONELLA PNEUMOPHILA*, teñida con acetato de uranilo y citrato de plomo y aumentada unas 50.000 veces en esta micrografía obtenida de un saco de huevo embrionado infectado. El organismo visto en secciones longitudinales y transversales es estructuralmente similar a otras bacterias, pero no está relacionada con otras especies y géneros descritos hasta su identificación; ello explica el porqué de su nombre de *Legionella pneumophila*.**

primer resultado. Una muestra de bazo teñida con la técnica de Giménez contenía pequeños organismos bacilares.

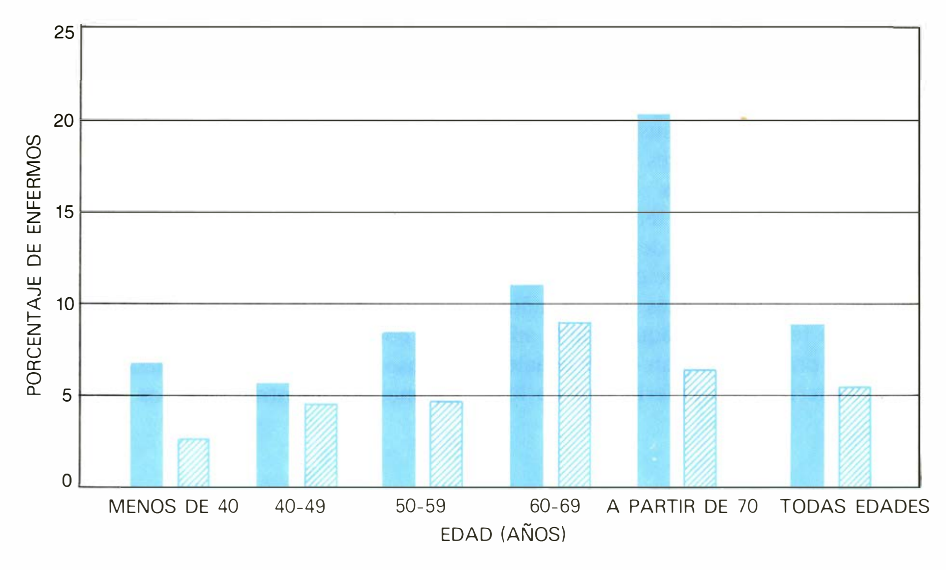
La observación en sí no era gran cosa. La mera presencia de unos cuantos organismos en el bazo del cobayo no nos permitía relacionarla con el organismo que había causado la muerte del enfermo, pero constituía una pista prometedora. Para intensificar el crecimiento de estos y otros organismos presentes en los tejidos animales, las suspensiones de bazo de cobayo se inocularon en el saco vitelino de huevos de gallina embrionados. Los embriones murieron días después. Se separaron las membranas del saco y se tiñeron por el método de Giménez algunos frotis de estas membranas. Pudo observarse un cultivo, aparentemente puro, de organismos bacilares; pero no logramos establecer si se trataba de grandes rickettsias o de bacterias de pequeño tamaño. E inmediatamente procedimos a investigar si era la causa de la enfermedad de los legionarios.



DISTRIBUCION de los casos en la epidemia de Filadelfia. La curva que podría dibujarse es la típica de un brote que tiene una única fuente u origen: el número de casos entre los asistentes a la convención (*color intenso*) y entre las personas ajenas a la misma (*color sombreado*) se elevó rápidamente a partir del 22 de julio hasta el 25 del mismo mes, descendiendo luego más lentamente. La relación entre las fechas de la convención militar norteamericana y el inicio de la enfermedad entre los legionarios indicaba que el período de incubación de la enfermedad de los legionarios fluctuó entre dos y diez días.

Establecimiento de la causa

Se llevó a cabo investigando la presencia, en el suero del paciente de la enfermedad de los legionarios, de anticuerpos específicos para el organismo recientemente aislado. Cuando el suero de un paciente contiene anticuerpos específicos para las moléculas antigénicas de un determinado organismo, recibe el nombre de "seropositivo"; en cuyo caso puede asumirse que el enfermo ha estado alguna vez expuesto a dicho organismo en particular (o posiblemente a uno con antígenos similares). Si se observa que el nivel de anticuerpos del enfermo se eleva en el curso de la convalecencia de una determinada enfermedad, queda demostrada la "seroconversión" y es muy probable que la enfermedad haya sido provocada por dicho organismo.



INCIDENCIA de la enfermedad epidémica de los legionarios, de 1976, entre los delegados de la convención de Filadelfia. Varió con respecto a la edad y el hotel en el que residieron. La incidencia fue mayor en adultos que en jóvenes, así como fue también mayor entre los delegados que se hospedaron en el hotel Bellevue-Stratford (*color intenso*) que en los que estuvieron en hoteles distintos (*sombreado*).

Al ensayar por inmunofluorescencia indirecta [*véase la ilustración de la página 54*] sueros procedentes de los enfermos frente a los anticuerpos específicos para el organismo recientemente aislado, más del 90 por ciento de las muestras de suero tomadas durante la convalecencia resultaron ser seropositivas, y en más del 50 por ciento de los casos –la mayoría de aquellos de los que se disponía de muestras de suero convenientemente espaciadas en el tiempo– pudimos demostrar seroconversión, lo que indicaba que los enfermos habían sido infectados recientemente por este particular bacilo.

Aunque los resultados de la inmunofluorescencia indirecta mostraban que habíamos aislado el organismo causal de la enfermedad de los legionarios, durante varias semanas no supimos si se trataba de una rickettsia o de una bacteria. Morfológicamente podía ser una bacteria, pero había sido aislada con técnicas más adecuadas para las rickettsias, y era incapaz de crecer en una extensa gama de medios de cultivo en los que normalmente las bacterias logran desarrollarse. Finalmente, Robert E. Weaver, del CDC, encontró un medio bacteriológico en el cual podía crecer este bacilo. Era similar a un medio en el que se aisló el gonococo. Se preparó por adición de un uno por ciento de hemoglobina y otro uno por ciento de un suplemento comercial (IsovitaleX) a un medio estándar (agar de Mueller-Hinton). Cuando Weaver inoculó este medio con una suspensión concentrada de material procedente de un saco vitelino infectado se observó, tras algunos días de incubación, que habían aparecido varias colonias bacterianas sobre el medio de cul-

tivo. El bacilo aislado no era pues una rickettsia sino una bacteria.

Se trataba de una bacteria extraordinariamente exigente, con muchos requerimientos nutritivos específicos. James C. Feeley, del CDC, junto con sus colaboradores delimitó esas exigencias. Entre otras cosas, la bacteria no se desarrollaba si el medio de cultivo no contenía una concentración suficientemente elevada del aminoácido cisteína (del que hay gran cantidad en el producto IsoVitaleX) así como de hierro (suministrado por la hemoglobina). Tales necesidades nutritivas son muy poco corrientes pues, aun cuando muchas bacterias requieren cisteína o hierro como suplemento, el doble requerimiento es raro.

Otra característica poco corriente de la bacteria de la enfermedad de los legionarios era que en las muestras patológicas (por ejemplo, tejido pulmonar de un enfermo) no se teñían bien con técnicas que suelen mostrarse eficaces para bacterias. Incluso después de comprobarse (gracias a nuestro éxito en el aislamiento del organismo en tejido pulmonar) que la bacteria estaba presente en los pulmones afectados, no podían observarse los organismos en las secciones de pulmón teñidas por métodos convencionales. Francis W. Chandler y sus colegas en el CDC ensayaron entonces una serie de métodos convencionales. Descubrieron al fin que la bacteria se teñía bien si recurrían a una variación de la impregnación con plata de Dieterle, procedimiento desarrollado más de 50 años antes para teñir la espiroqueta de la sífilis. Estas propiedades tan poco usuales de la bacteria de la enfermedad de los legionarios sugerían que se trataba de una especie desconocida, pero había que caracterizarla con mayor precisión antes de emitir un juicio definitivo.

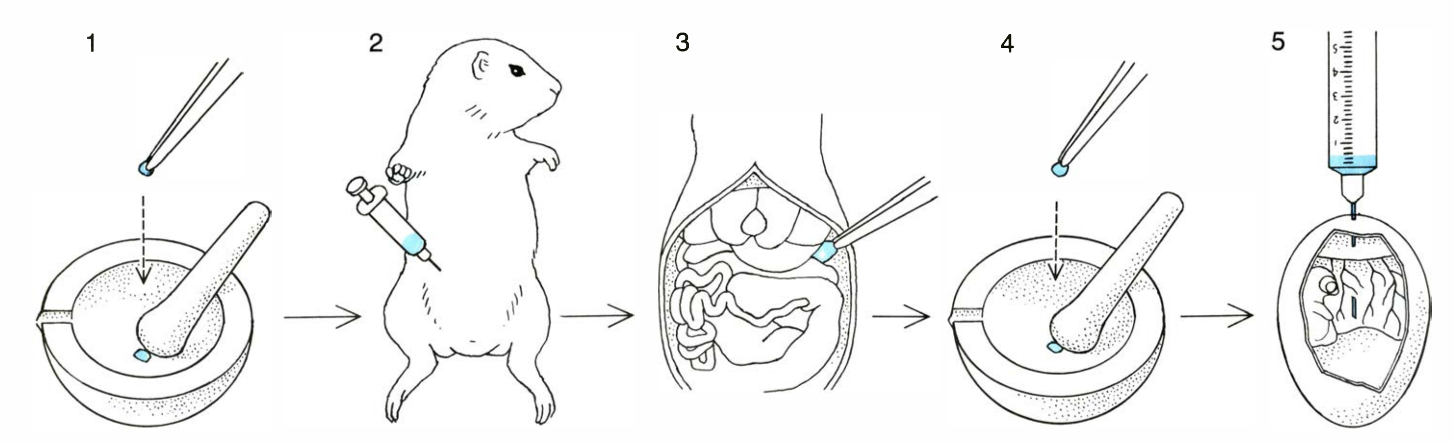
Brotes anteriores

Paralelamente, otra línea de estudio que se estaba llevando a cabo dentro del programa de investigación intensiva del CDC comprobó que no se trataba de una nueva bacteria, sino de un especimen con el que los investigadores no estaban familiarizados. Una cuestión importante planteada a raíz del descubrimiento de este agente fue determinar si este bacilo poco conocido era o no el responsable de brotes anteriores de enfermedades respiratorias cuya etiología seguía ignorándose. Quizás el más espectacular de ellos fuera la "fiebre Pontiac" de 1968. Apareció el brote el 2 de julio de ese año. El 95 por ciento de las personas que trabajaban en un mismo edificio del Departamento de Sanidad del condado de Oakland, en Pontiac (Michigan), padecieron fiebre, jaqueca y dolores musculares; otros síntomas comunes a los afectados fueron diarrea, vómitos y dolor pectoral, aunque ninguno llegó a enfermar de neumonía. La afección duró sólo tres o cuatro días, y todos los enfermos se recobraron. Un equipo de epidemiólogos del CDC, encabezados por Thomas H. Glick y Michael B. Gregg, investigaron el brote, enfermando algunos de ellos uno o dos días después de haber entrado en el edificio. Los cobayos que Arnold F. Kaufmann expuso al aire del edificio desarrollaron neumonía, mientras que los animales control situados en edificios adyacentes permanecieron sanos.

Se observó que, de entre las personas que entraron en el edificio, las únicas que no se vieron afectadas eran las que se encontraban en él cuando el aire acondicionado estaba desconectado. La consecuente revisión del sistema de aire acondicionado permitió descubrir un defecto en su funcionamiento. El vapor originado por el condensador de evaporación no quedaba eliminado convenientemente; se descubrió la existencia de unos orificios en el respiradero de aire consumido que conectaban con el respiradero de aire fresco, lo que permitió que el vapor se condensara en un recodo del conducto de aire fresco. Algunos cobayos fueron entonces rociados con agua procedente de este recodo y enfermeron de neumonía. La conclusión a la que se llegó entonces fue que el agente de la fiebre Pontiac se encontraba en el agua del condensador, diseminada a través del sistema de aire acondicionado. Sin embargo, no se pudo aislar ningún organismo patógeno, ni se supo de la causa de la enfermedad.

En el CDC se habían conservado muestras de suero de enfermos de fiebre Pontiac. En 1977, nueve años después, se ensayaron frente a anticuerpos del recién aislado bacilo de la legionelosis. De suero procedente de 37 casos, 31 muestras evidenciaron seroconversión al organismo durante la convalecencia, lo que señalaba que la fiebre Pontiac había sido originada por el agente que provocó la enfermedad de los legionarios.

Otro brote anterior que ha sido atribuido al mismo organismo se produjo



AISLAMIENTO de la bacteria *L. pneumophila*. Se consiguió por vez primera gracias al empleo de un método usual en el aislamiento de microorganismos no bacterianos, tales como las rickettsias. Se trituró un fragmento de tejido pulmonar de un enfermo (*1*), se suspendió en una solución tampón y se inoculó en la cavidad abdominal de cobayos (*2*). Los animales enfermaron. En la autopsia (*3*), un fragmento de bazo que presumiblemente contenía el todavía desconocido agente patógeno se colocó en una suspensión (*4*), que se inoculó en sacos de huevos embrionados (*5*) para intensificar el crecimiento de los organismos que pudieran estar presentes. Los embriones murieron a los cuatro o cinco días. Se extrajo la membrana del saco vitelino de los embriones muertos (*6*), que se extendió sobre un portaobjetos (*7*) y se coloreó por el método de Giménez (*8*). En el microscopio pudieron observarse unos organismos alargados (*9*). A fin de obtener un cultivo del organismo se hizo una suspensión de una membrana embrionaria (*10*), con la cual se sembraron por estría varios medios bacteriológicos (*11*). Las colonias pudieron desarrollarse en un agar de Mueller-Hinton enriquecido (*12*), poniendo de manifiesto que se trataba de una bacteria y a la vez permitiendo su utilización en varios experimentos posteriores. Se han diseñado medios suplementados que satisfacen los requerimientos específicos de este organismo, gracias a los cuales la bacteria puede aislarse ya directamente.

entre julio y agosto de 1965, cuando 81 enfermos del Hospital de Sta. Isabel, un centro reservado a enfermos crónicos en Washington (distrito de Columbia), padecieron neumonía, muriendo catorce de ellos. La causa, entonces, no pudo determinarse. En 1977 se contrastaron muestras de suero de fase aguda y de convalecencia de 23 enfermos con el agente productor de la enfermedad de los legionarios. En 17 de los enfermos se observó un incremento importante de anticuerpos contra esta bacteria. Mediante ensayos similares se pudo relacionar la bacteria con varios brotes anteriores, entre ellos uno que había afectado en 1973 a un grupo de escoceses que pasaba sus vacaciones en España.

La publicación de un trabajo sobre las características de la bacteria de la enfermedad de los legionarios también llevó a la conclusión de que se trataba del mismo agente responsable de otros brotes anteriores. F. Marilyn Bozeman, de la U. S. Food and Drug Administration, informó que la bacteria era similar en algunos aspectos a cuatro agentes infecciosos parecidos a las rickettsias que ella había descrito años antes, y que aún no habían podido ser clasificados. Al igual que la nueva bacteria aislada, habían sido aislados a partir de huevos embrionados infectados con tejidos de cobayos inoculados, bien con muestras clínicas, bien con muestras de autopsia, que se teñían también por el método de Giménez. Durante las investigaciones de Bozeman, los agentes no identificados se habían almacenado en un ultracongelador. En 1978 diversos ensayos revelaron que tres de los cuatro agentes similares a las rickettsias no estaban relacionados con el agente causal de la enfermedad de los legionarios, pero que el cuarto era virtualmente idéntico. La bacteria idéntica había sido aislada en 1947 a partir de un cobayo inoculado con sangre procedente de un enfermo que padecía una enfermedad respiratoria que producía fiebre.

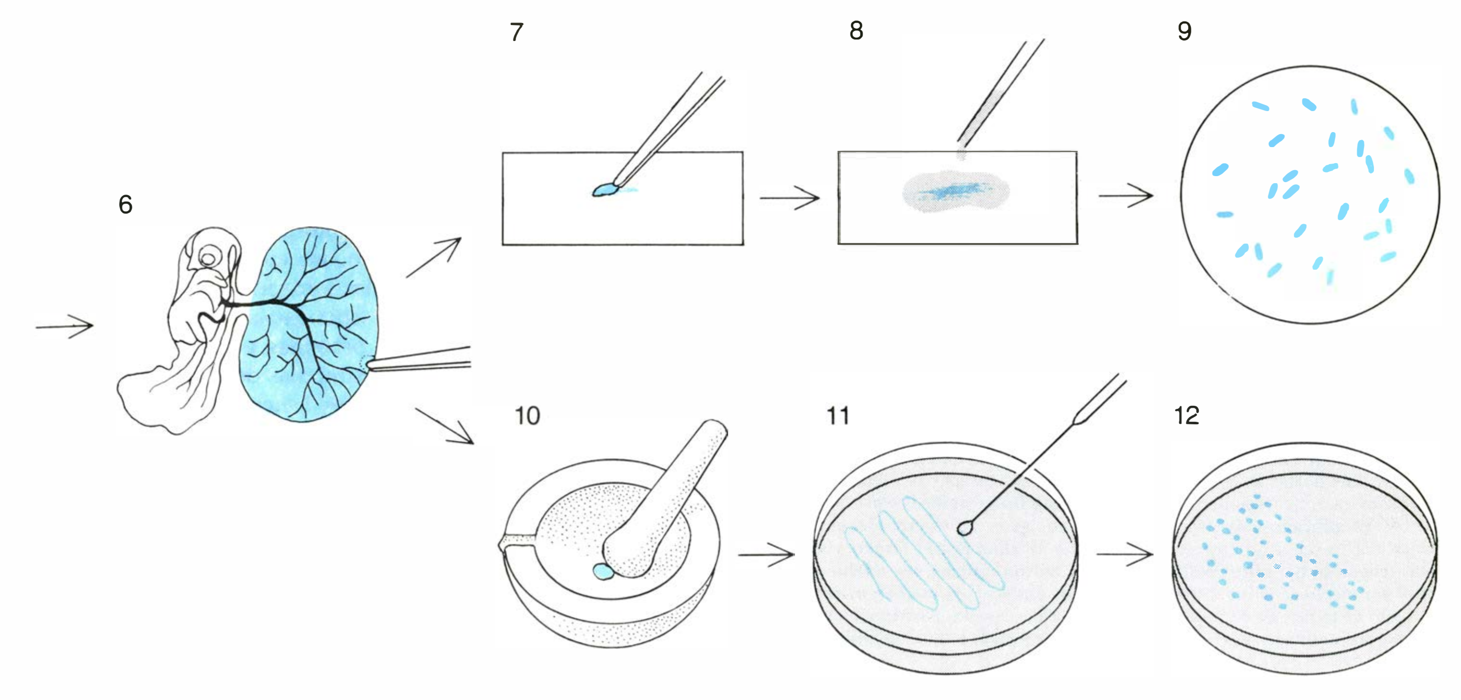
Caracterización de la bacteria

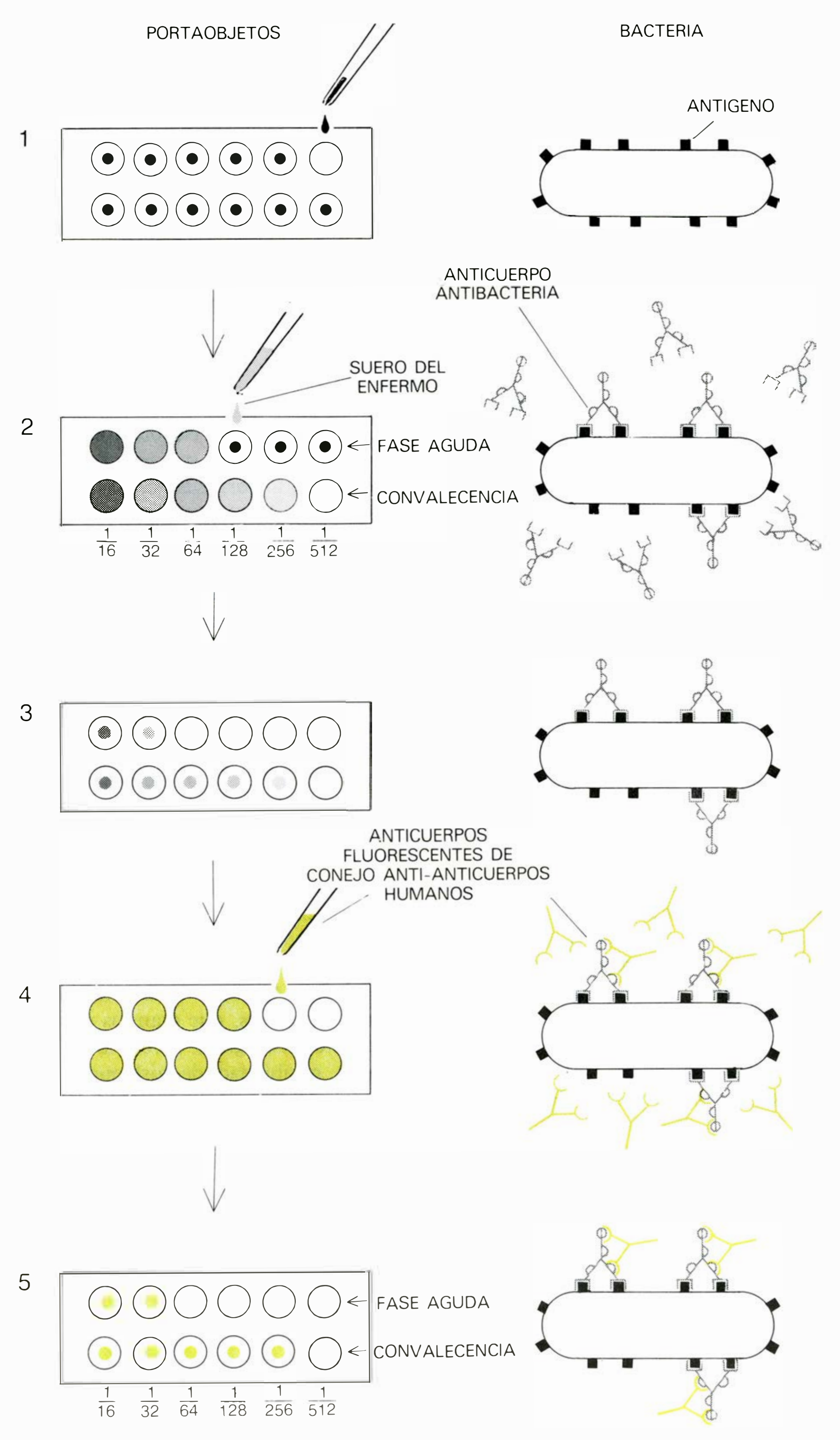
No cesaron aquí los esfuerzos por conocer mejor la bacteria de la enfermedad de los legionarios. Para asegurar que una bacteria pertenece a una nueva especie debe demostrarse, de forma convincente, que es significativamente distinta de todas las especies hasta la fecha descritas. Un equipo de microbiólogos del CDC examinaron la morfología, fisiología y propiedades tintoriales del nuevo microorganismo, su susceptibilidad a varios antibióticos, su dotación antigénica y su composición de ácidos grasos. Se compararon esas propiedades con las de las bacterias conocidas. De ello resultó que el agente de la enfermedad de los legionarios compartía ciertas características con otras bacterias, pero el conjunto de sus propiedades era muy distinto de las del resto de especies.

La prueba definitiva de que este organismo era una nueva especie se obtuvo cuando Donald J. Brenner, Arnold G. Steigerwalt y sus colegas, del CDC, compararon el material genético de la bacteria de Filadelfia con el de otras bacterias por medio de la depurada técnica de la hibridación de su ADN. El cromosoma de la bacteria de la enfermedad de los legionarios, al igual que el de otras bacterias, es una molécula de ADN de doble hebra compuesta de dos cadenas simples unidas entre sí gracias al apareamiento de sus nucleótidos complementarios. Bajo condiciones experimentales apropiadas, los filamentos de ADN de especies afines pueden unirse para formar moléculas híbridas de doble cadena; el grado de hibridación depende de la mayor o menor proximidad genética [*véase la ilustración de la página 55*].

Los intentos de hibridar cadenas simples de ADN de la bacteria de la enfermedad de los legionarios con filamentos de otras bacterias no permitieron identificar ninguna especie que estuviera relacionada con los organismos aislados recientemente. Aun cuando todavía no se ha cerrado el cupo de comparaciones, la prueba experimental disponible parece justificar la designación de la bacteria como un nuevo género y especie. Y, aunque de un modo provisional, ya ha sido bautizada: *Legionella pneumophila*.

Si echamos la vista atrás, es fácil comprender por qué *L. pneumophila* no se había detectado, a pesar de haber causado enfermedades durante décadas. En primer lugar, la enfermedad que origina es clínicamente similar a ciertos tipos de neumonía no bacteriana ("atípica"), y en la investigación del agente causal de tales enfermedades los investigadores solían incorporar antibióticos en las muestras para reducir la contaminación bacteriana. Esto facilitaba el aislamiento de virus o rickettsias, que eran los supuestos agentes etiológicos, pero a la vez excluía la posibilidad de aislamiento de bacterias del tipo de *L. pneumophila*. La





ENSAYO DE INMUNOFLUORESCENCIA INDIRECTA, que permitió establecer que la bacteria aislada había provocado la enfermedad de los legionarios. Se describe aquí a nivel experimental (*izquierda*) y a nivel del organismo (*derecha*). Se colocó una gota de una suspensión de membrana embrionaria infectada en cada uno de los doce pocillos de un porta (*1*); se fijó con acetona. En cada una de las dos series, se introdujeron seis diluciones sucesivas de suero tomado de un enfermo en fase aguda y de otro en convalecencia (*2*); todos los anticuerpos específicos (*estructuras grises en Y*) presentes en el suero se fijaban a los antígenos a nivel de la membrana externa durante un período de incubación de 30 minutos. Los portas se lavaron eliminando los anticuerpos no fijados y otros componentes séricos (*3*). Se añadió un anticuerpo de conejo marcado con fluoresceína (*color*), específico para anticuerpos humanos (*4*). El anticuerpo de conejo se unía a cualquier anticuerpo del enfermo y permanecía fijado a la bacteria. El exceso de anticuerpo de conejo se eliminó por lavado y cada pocillo se examinó en un microscopio de luz ultravioleta. Las bacterias con anticuerpos unidos emitían un brillo verde amarillento por efecto de la luz ultravioleta (*5*). Los anticuerpos estaban presentes en diluciones más elevadas del suero de la fase de convalecencia que de la fase aguda, mostrando que el enfermo había desarrollado anticuerpos antibacteria en el curso de la enfermedad, lo que identificó a la bacteria como el causante.

bacteria pudo al fin aislarse en el curso de un protocolo de aislamiento de rickettsias en el que no se incluyeron antibióticos. En segundo lugar, los peculiares requerimientos nutritivos de la bacteria obstruyeron su identificación precoz. Los medios en los que se aislan los hongos y bacterias que usualmente provocan neumonía no contienen las concentraciones adecuadas de cisteína y hierro. Por último, la imposibilidad de visualizar correctamente la bacteria en tejidos teñidos por los métodos acostumbrados debe haber sido ciertamente uno de los principales obstáculos con que toparon las investigaciones de los primeros casos de neumonía atípica causada por *L. pneumophila*.

Curso clínico y terapia

La legionelosis, así se llama ahora técnicamente la enfermedad causada por *L. pneumophila*, se ha observado hasta el presente en dos formas básicas: la enfermedad de los legionarios y la fiebre Pontiac. Los rasgos que distinguen una de la otra son: el período de incubación, la incidencia de la enfermedad y la presentación de neumonía.

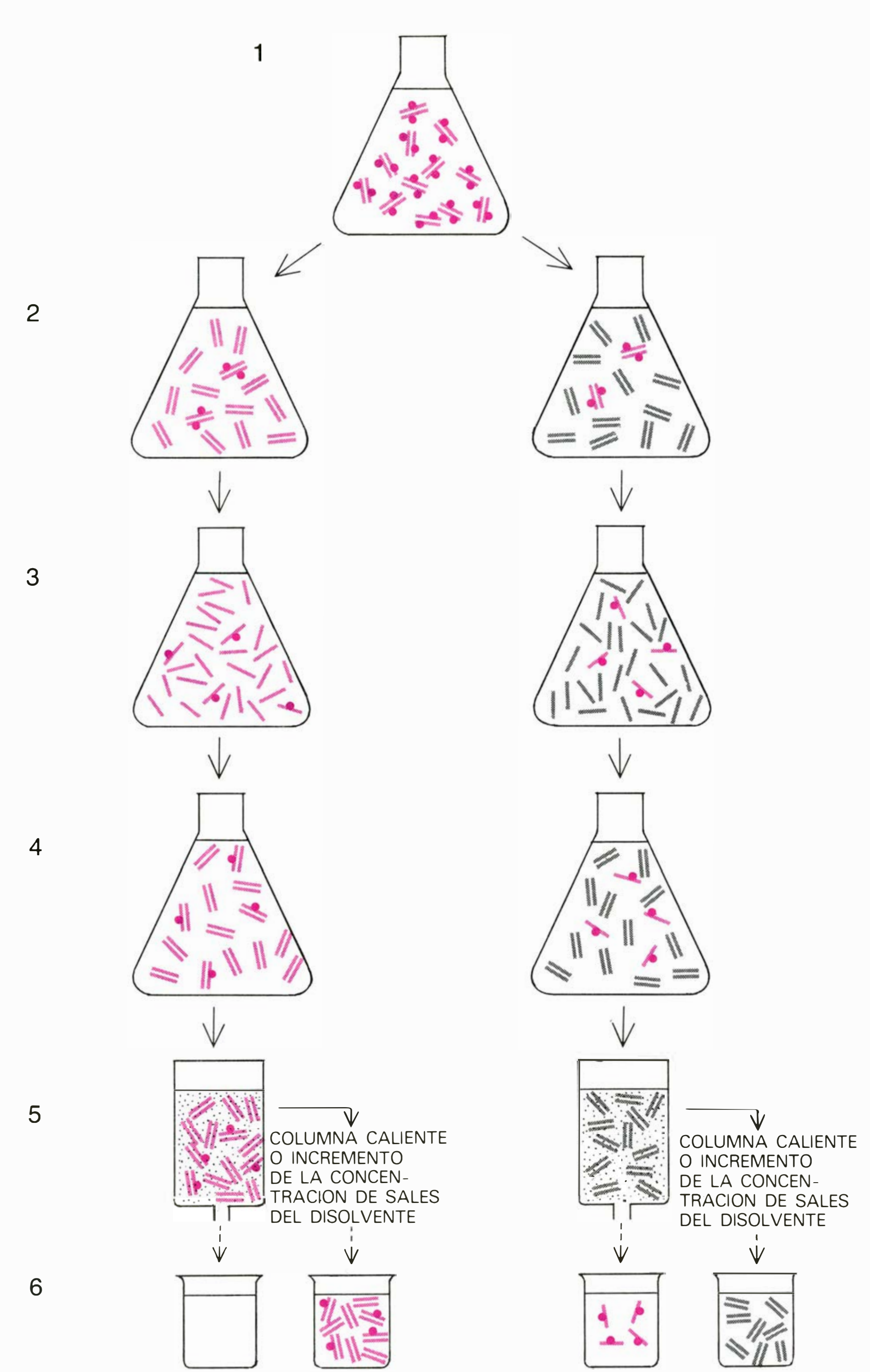
La enfermedad de los legionarios se manifiesta típicamente entre los dos y los diez días después de la infección. Un malestar general va asociado a dolores musculares y un ligero dolor de cabeza. La fiebre aumenta rápidamente y con frecuencia se acompaña de tos, dolores de tórax y abdomen, diarrea y aumento del ritmo respiratorio (disnea). Al ser reconocidos por un médico, la mayoría de enfermos están a una temperatura de 39 o 40 grados Celsius, y sufren estertores pulmonares (sonidos respiratorios anormales que se detectan por medio de un estetoscopio y que sugieren la existencia de una infección pulmonar). Algunos enfermos aparecen desconcertados y como atontados, lo que da a entender que el sistema nervioso central resulta afectado. Los análisis muestran generalmente un recuento de leucocitos moderadamente alto, un incremento del número de células inmaduras, una elevada velocidad de sedimentación de hematíes, presencia de proteínas y glóbulos rojos en la orina y síntomas de algunas anormalidades en las funciones hepática y renal. En el 90 por ciento de los enfermos las radiografías de tórax revelan neumonía.

Los enfermos afectos de la enfermedad de los legionarios requieren en general hospitalización. Durante varios días después del ingreso, sus condiciones van empeorando. En ausencia de un tra-

tamiento específico, cerca del 20 por ciento de estos enfermos mueren por neumonía progresiva (que interfiere en la oxigenación de la sangre) o por shock (cuyo mecanismo se desconoce pero que puede resultar de la presencia del microorganismo en la sangre). El 80 por ciento restante se recupera gradualmente al cabo de una semana, o algo más, durante la cual algunos pacientes precisan de ayuda mecánica en la respiración o de diálisis para tratar el fallo renal. La recuperación suele ser total, aunque los enfermos puedan sentirse débiles durante algunos meses y en cierta medida los pulmones queden afectados permanentemente.

Como la enfermedad de los legionarios, la fiebre Pontiac se inicia con dolores de cabeza, dolores musculares y fiebre; típicamente, los síntomas hacen su aparición al día, o a los dos días, después de la infección. Los síntomas son tos, dolor torácico, diarrea, vómitos, dolores de garganta y confusión, pero normalmente el cuadro clínico no evoca nada espectacular. En la fiebre Pontiac no se presenta neumonía ni shock, ni tampoco resultan afectados el riñón y el hígado. Los enfermos presentan fiebre y una gran debilidad durante unos dos o cinco días, si bien todos se recuperan, al menos aparentemente, de forma completa.

Sigue sin saberse cuáles son los antibióticos más eficaces contra las infecciones de *L. pneumophila*. No se han hecho todavía pruebas al azar con controles en seres humanos. Claire V. Broome y sus colegas, del CDC, observaron que durante una epidemia registrada en Vermont en 1977 sólo una persona de las 22 que habían sido tratadas con eritromicina murió, mientras que de 41 que no habían recibido eritromicina fallecieron 10. Clyde Thornscerry descubrió que se inhibiá el crecimiento de la bacteria en agar por concentraciones relativamente bajas de varios antibióticos, pero el organismo fabricaba un enzima que inactivaba a la mayoría de antibióticos de un grupo determinado, las cefalosporinas. Los embriones de pollo infectados podían protegerse bien con eritromicina, rifampicina, gentamicina, estreptomicina, sulfadiacina o cloramfenicol. Con Theodore F. Tsai y otros colegas, observamos que los cobayos infectados se recuperaban tras un tratamiento con eritromicina o rifampicina. Ahora, la ciencia recomienda a los médicos tratar a los pacientes que presentan la enfermedad de los legionarios con eritromicina, añadiendo la terapia de rifampicina en aquellos casos en que no se ob-



ENSAYO DE HIBRIDACION DEL ADN. El test se basa en que las dos cadenas de la molécula de ADN de doble filamento están unidas por enlaces entre nucleótidos complementarios y específicos; en consecuencia los filamentos aislados de la misma especie vuelven a aparearse fácilmente y los procedentes de especies distintas lo hacen en proporción a la proximidad existente entre ellas. Se marca una doble cadena de ADN de *L. pneumophila* (*color claro*) con un isótopo radiactivo (*punto coloreado*) y se rompe en fragmentos cortos (*1*). Una pequeña cantidad de ADN marcado se mezcla con fragmentos cortos de ADN no marcado (*2*) de la misma bacteria (*izquierda*) y de otra bacteria distinta (*gris*) de la que quiere ensayarse su proximidad genética (*derecha*). Las mezclas se calientan para romper los enlaces entre los nucleótidos complementarios y obtener filamentos simples de ADN (*3*). Al enfriar, y mezclar luego, las mezclas, las cadenas complementarias se aparean (*4*). Prácticamente todas las cadenas de ADN mezcladas con ADN homólogo "encuentran" una cadena complementaria no marcada (*izquierda*), pero no así los fragmentos de ADN marcados mezclados con ADN de una bacteria distinta (*derecha*). El ADN marcado está presente en una concentración tan baja que la reasociación al azar entre cadenas marcadas es despreciable; todo el ADN de doble cadena marcado es un híbrido de cadenas marcadas y no marcadas. El ADN de doble cadena y el de cadena sencilla se separan pasándolos por una columna que contiene gel de hidroxilapatita (*5*); se extraen (*6*). Midiendo las cantidades relativas de radiactividad en los ADN sencillo y doble se tiene una estima del nivel de hibridación. En este esquema no hay hibridación entre el ADN de *Legionella pneumophila* y el ADN de la bacteria frente a la que se ha ensayado, lo que demuestra ausencia de proximidad genética entre ambas.

serve respuesta. En principio la terapia con antibióticos (antibioterapia) no es necesaria en la fiebre Pontiac.

Diagnóstico

El cuadro clínico que configura la enfermedad de los legionarios y sirve para distinguirla de otras neumonías registra: diarrea, confusión, respiración forzada, dolor de garganta, glóbulos rojos en la orina, disfunciones hepáticas, ausencia en el esputo de bacterias típicamente productoras de neumonía y, evidentemente, falta de respuesta a los tratamientos que han resultado ser inútiles frente a *L. pneumophila*.

Sólo haremos un diagnóstico específico de la legionelosis mediante aislamiento del organismo a partir del enfermo, por comprobación de un incremento en el nivel de anticuerpos específicos cuando el paciente se recupera, o demostrando la presencia de *L. pneumophila* en los tejidos o en los fluidos orgánicos del enfermo. El organismo se ha aislado a partir de sangre, esputos, líquido pleural y tejidos pulmonares por inoculación de cobayos o de medios de cultivo bacteriológicos adecuados. Sin embargo, existen ciertas limitaciones prácticas que dificultan el aislamiento de la bacteria directamente sobre un medio de cultivo. Puede resultar difícil obtener muestras bacterianas adecuadas, toda vez que ese organismo no se cultiva con facilidad y transcurren varios días hasta que el cultivo crece.

Por medio de la técnica de inmunofluorescencia indirecta, al final de la segunda semana de enfermedad pueden apreciarse elevadas concentraciones de anticuerpos en la mitad casi de los enfermos estudiados, aunque algunos pacientes no fabrican anticuerpos hasta la sexta semana y cerca del 15 por ciento de estos últimos no alcanzan nunca niveles significativos. Existen otras técnicas para la detección de anticuerpos, pero no se ha comprobado suficientemente su sensibilidad y especificidad. Otro problema es la existencia de varios grupos serológicos en *L. pneumophila*, cada uno de los cuales incorpora un grupo distinto de antígenos. Cualquier miembro de uno de los grupos puede desencadenar la infección; por tanto, habremos de ensayar el suero del enfermo frente a anticuerpos representativos de cada grupo. Todavía más: un diagnóstico basado en un incremento del nivel de anticuerpos específicos será necesariamente retrospectivo. Sólo durante la convalecencia se admite que el nivel de anticuerpos sea suficientemente elevado como para confirmar un diagnóstico, lo que obviamente resulta demasiado tarde para servirle de ayuda al médico a la hora de aplicar el tratamiento idóneo.

El ensayo por inmunofluorescencia directa, en cambio, sí permite el diagnóstico en unas pocas horas, en los comienzos del proceso infeccioso. Los esputos de enfermos u otras secreciones respiratorias se exponen a anticuerpos anti-*Legionella pneumophila* (preparados por inmunización de conejos con bacterias inactivadas) que previamente se han combinado con un material fluorescente. Los anticuerpos se fijan a los antígenos de *L. pnemophila* poniendo de manifiesto la presencia del microorganismo. Un 70 por ciento de los enfermos resultan positivos y los errores en la apreciación parecen ser poco frecuentes. Los trabajos preliminares sugieren que los antígenos del microorganismo pueden detectarse en suero y orina de enfermos por un complejo procedimiento denominado reacción ELISA (del inglés "enzyme-linked immunosorbence"). Si esta prueba demuestra ser suficientemente sensible y específica, podría resultar de gran valor en los diagnósticos precoces.

	ENFERMEDAD LEGIONARIOS	FIEBRE PONTIAC
INCIDENCIA	1-5 POR CIENTO	95 POR CIENTO
PERIODO DE INCUBACION	DE 2 A 10 DIAS	UNO O DOS DIAS
SINTOMAS	FIEBRE, TOS, DOLOR MUSCULAR, ESCALOFRIOS, DOLOR DE CABEZA, DOLOR TORACICO, ESPUTOS, DIARREA, CONFUSION	FIEBRE, TOS, DOLOR MUSCULAR, ESCALOFRIOS, DOLOR DE CABEZA, DOLOR TORACICO, CONFUSION
EFECTOS EN PULMON	NEUMONIA, EFUSION PLEURAL	PLEURITIS, AUSENCIA DE NEUMONIA
AFECCION EN OTROS ORGANOS	RIÑON, HIGADO, TRACTO GASTROINTESTINAL, SISTEMA NERVIOSO	NINGUNO
PROPORCION DE CASOS FATALES	15-20 POR CIENTO	AUSENCIA

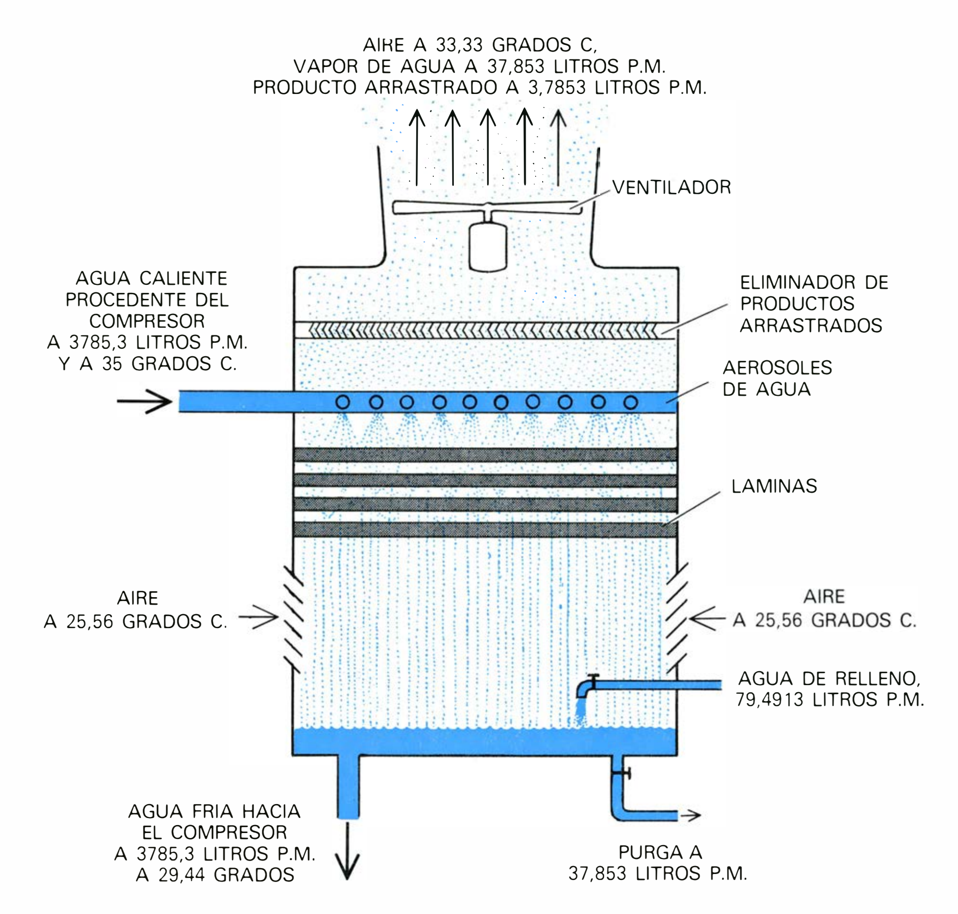
LEGIONELOSIS se denomina la enfermedad provocada por *L. pneumophila*. Se ha observado básicamente en dos formas: la enfermedad de los legionarios y la fiebre Pontiac. Los rasgos distintivos principales son la incidencia (mucho más elevada en la fiebre Pontiac), el período de incubación (más corto en la fiebre Pontiac) y la presencia de neumonía en la enfermedad de los legionarios, pero no en la fiebre Pontiac, afección ésta mucho más moderada, y cuyo brote epidémico se detectó en el año 1968.

Incidencia de la enfermedad

En los cortos años transcurridos desde que se reconoció la legionelosis como una nueva enfermedad, se han descubierto casos prácticamente dondequiera que se han buscado: en más de 40 estados de la Unión, en Australia, Canadá, Dinamarca, Grecia, Israel, Italia, Holanda, Noruega, España, Suecia, Suiza, Alemania Federal, Inglaterra y Yugoslavia. No ha aparecido de momento un patrón geográfico, aunque la información procedente de muchas zonas es aún escasa o incompleta. A veces existen grupos de casos que parecen tener un origen común y en ocasiones se dan episodios aislados. Casi todos los casos del grupo y la mayoría de los esporádicos se han detectado en la estación de verano.

La enfermedad de los legionarios se presenta en individuos de mediana edad o ancianos, aunque se han dado algunos casos en niños de 16 meses. Se ha registrado mayor incidencia en varones que en mujeres. La enfermedad es más común en fumadores de cigarrillos (pero no en los que prefieren cigarros puros o pipa) y en grandes bebedores. Problemas clínicos subyacentes, como el cáncer, leucemia, linfoma y fallo renal, así como los fármacos que debilitan el sistema inmunitario, parecen predisponer al individuo para la enfermedad de los legionarios (y para otras infecciones).

Gregory A. Storch y William B. Baine, del CDC, han encontrado que los casos esporádicos se encuentran, sobre todo, entre viajeros, trabajadores de la construcción y personas que viven cerca de lugares de excavación o construcción. Es posible que estos individuos se expongan más fácilmente al polvo arrastrado por el aire que transporte la bacteria. Los brotes de legionelosis se han dado, en general, dentro o cerca de un edificio, frecuentemente un hotel o un hospital. En algunas ocasiones, el edificio parece haber sido sencillamente un lugar al que ha concurrido gran número de individuos susceptibles, pero en otras el sistema de aire acondicionado central



LAS TORRES DE REFRIGERACION de los sistemas de aire acondicionado están implicadas en la propagación de *L. pneumophila*. Una torre de refrigeración normal de un tamaño apropiado para un hotel, hospital o un edificio comercial, puede reciclar unos 3785,3 litros de agua por minuto. El agua refrigerante procedente del compresor se pulveriza sobre las láminas, un ventilador impulsa el aire fresco por entre el aerosol, para optimizar la evaporación, por efecto de la cual se enfría el agua. Parte del agua se elimina en forma de vapor y otra pequeña cantidad en forma de corriente (gotitas de agua): el resto se recicla. Las bacterias transportadas por el aire podrían ser impulsadas por el ventilador, introducidas en el aerosol de agua y diseminadas en la corriente. Se descubrió un caso en el que las gotas proyectadas desde la torre de refrigeración contaminaron el aire acondicionado del inmueble.

estaba involucrado en la dispersión de la bacteria (por razones que quedarán aclaradas más adelante).

Ecología de la enfermedad

La incidencia global de la legionelosis no se conoce con precisón. Cuando Hjordis M. Foy, de la Escuela de salud pública de la Universidad de Washington, y sus colaboradores hicieron un estudio prospectivo de neumonías entre una población incluida en un programa sanitario, sus resultados indicaron que *L. pneumophila* era la causa de 12 casos por cada 100.000 habitantes y año. Si estas tasas se aplican a nivel nacional, se obtiene una cifra anual de unos 26.000 casos de legionelosis en los Estados Unidos. Entre el 0,3 y el 1,5 por ciento de las muestras séricas de enfermos de neumonía remitidas al CDC han resultado ser positivas a la legionelosis. Si estas muestras son en verdad representativas, suele considerarse que de los 2,4 millones de neumonía registradas al año, entre 7000 y 36.000 constituyen, en realidad, episodios de enfermedad de los legionarios.

Muchos agentes infecciosos que provocan epidemias de neumonía, pero no se propagan de persona a persona, lo hacen a través del aire a partir de un nicho ecológico característico en un medio ambiente no humano. Un ejemplo común es el hongo *Histoplasma capsulatum*. Vive en el suelo y provoca la enfermedad cuando se inhala el polvillo que él contamina. Parecía, pues, probable que *L. pneumophila* se desenvolviera en un ambiente inorgánico. Esta hipótesis se confirmó con estudios de laboratorio que mostraban que la bacteria podía sobrevivir (aunque aparentemente sin proliferar) durante más de un año en agua corriente.

El brote de 1965, localizado en el Hospital de Sta. Isabel, sugirió que *L. pneumophila* podía vivir en el suelo. En aquel verano, se había excavado el terreno en varios puntos del centro para la instalación de un nuevo sistema de riego del césped. Stephen B. Thacker y John V. Bennett, del CDC, descubrieron que eran principalmente los pacientes cuyas camas se hallaban próximas a las ventanas y en los edificios cercanos a la excavación quienes contrajeron la enfermedad. Los casos se presentaban agrupados en el tiempo, apareciendo un brote a los cinco o seis días después de cada excavación, lo que sugería que el polvo contaminado que se levantaba en el proceso se había diseminado hasta infectar a los enfermos a través del aire. En 1965 nadie pensó en buscar un agente infeccioso en el suelo (y, como hemos visto, de todas formas habría sido difícil aislar *L. pneumophila* de tales muestras), pero posteriormente el organismo ha podido ser aislado en el barro en otros lugares.

En tres de los brotes registrados se hallaba implicada, como fuente de infección, el agua de las torres de refrigeración o de los condensadores de evaporación de sistemas de aire acondicionado. En una torre refrigeradora el agua, calentada mediante un compresor, se pulveriza sobre varillas de madera u otro material, a medida que el aire se barre por un ventilador, con las gotitas que caen. Parte del agua se evapora, enfriándose el resto que pasa a reciclarse. Una pequeña cantidad de agua es arrastrada, cual chorro de corriente, en forma de vapor o de diminutas gotas con el aire consumido; la corriente arrastra con todo lo que hay en el agua de la torre de refrigeración, incluídas las bacterias presentes, y lo transporta hasta una considerable distancia. En un condensador de evaporación, el agua se pulveriza sobre espirales metálicas que contienen el refrigerante. Este dispositivo se enfría directamente por efecto de la evaporación del agua.

En dos brotes asociados a condensadores de evaporación y uno a torres de refrigeración, se comprobó que los enfermos habían estado expuestos a esa corriente. En las tres ocasiones se aisló *L. pneumophila* a partir del agua concentrada en el sistema de refrigeración. George K. Morris, del CDC, también aisló el microorganismo en torres de las que no se sabía que hubieran estado relacionadas con casos de legionelosis. Parece verosímil, pues, que torres y condensadores estén en general contaminados con bacterias procedentes del suelo, diseminadas por el aire y acumuladas por el agua de condensación. Si esto es cierto, quedan por definir las condiciones bajo las cuales torres y condensadores contaminados pueden llegar a diseminar las bacterias. Mientras tanto, se están realizando estudios a fin de encon-

trar agentes antibacterianos que sean eficaces frente a *Legionella pneumophila* y que sirvan para limpiar, de polutantes malignos, los equipos modernos de aire acondicionado.

Cuestiones por resolver

Parece paradójico que un organismo tan exigente en sus requerimientos nutritivos, y por tanto difícil de cultivar en el laboratorio, llegue a mantenerse con éxito en un medio inorgánico. Falta por establecer de un modo absoluto el nicho ecológico preciso de *L. pneumophila* y sus fuentes nutritivas. El desentrañamiento de su historia natural constituiría un gran paso hacia la comprensión de las circunstancias que dan lugar a las infecciones esporádicas aisladas y las condiciones de exposición a la bacteria que originan brotes epidémicos.

Todavía no se sabe por qué *L. pneumophila* provoca dos tipos distintos de enfermedades tales como la legionelosis y la fiebre Pontiac, ni por qué la bacteria afecta sólo al uno por ciento de la población aparentemente expuesta a la enfermedad de los legionarios y al 95 por ciento de los susceptibles de contraer la segunda. Podría ser que las cepas bacterianas presentaran diferencias en algunas características fundamentales que determinaran la enfermedad que desarrollan. Sin embargo, los ensayos de laboratorio, incluida una prospección de toxinas producidas por la bacteria, no han demostrado todavía que existan diferencias significativas entre la cepa Pontiac y la de Filadelfia.

Es muy posible que la enfermedad de los legionarios y la fiebre Pontiac no sean los únicos síndromes provocados por *L. pneumophila*. Los investigadores clínicos han empezado a estudiar recientemente la posibilidad de que la bacteria pueda ser responsable, por ejemplo, de ciertos casos de abscesos pulmonares y endocarditis (inflamación del revestimiento interno del corazón). Es más, sólo unos dos tercios de los 2,4 millones de casos de neumonías que se producen anualmente en los Estados Unidos pueden relacionarse con un agente vírico o bacteriano conocido; se desconoce la causa de los 800.000 casos restantes. La ciencia no ha ganado todavía el reto de la identificación de estos agentes. Quizás algunos de los agentes desconocidos sean bacterias relacionadas con *Legionella pneumophila*, que no se han detectado aún. Proseguir la investigación sobre la legionelosis podría constituir la senda para llegar al descubrimiento de esos organismos.

La membrana fotosintética

La conversión de energía lumínica en energía química, realizada por las plantas verdes, se lleva a cabo en la membrana tilacoidea de la célula vegetal. La microscopía electrónica revela la asimetría que hace posible tal conversión

Kenneth R. Miller

La Tierra es un planeta bañado en luz. No debiera sorprender, pues, que muchos de los organismos vivos que han evolucionado sobre ella hayan desarrollado la capacidad de captar energía lumínica. De entre todas las interacciones que se producen entre la vida y la luz, la principal de ellas es la fotosíntesis: la conversión biológica de energía lumínica en energía química. Desde un punto de vista energético, todo lo viviente depende en última instancia de la fotosíntesis, fuente de todas las formas de alimento, e incluso del oxígeno de la atmósfera terrestre. La mayoría de células, desde la simple alga hasta las grandes y complejas plantas terrestres, son fotosintéticas. Ultimamente, se ha avanzado mucho en el conocimiento del diseño y la función de la estructura biológica en la que tiene lugar la captación de energía lumínica: la membrana fotosintética. La reacción química global de la fotosíntesis puede expresarse por medio de una ecuación sorprendentemente sencilla: se toman seis moléculas de dióxido de carbono del medio ambiente junto con seis moléculas de agua y, en presencia de luz, estas moléculas se transforman en una molécula de azúcar de seis carbonos, la glucosa, liberándose a la vez seis moléculas de oxígeno. Los productos de la reacción tienen mayor energía química que los reactivos y así, en cierto sentido, puede decirse que la energía de la luz solar ha quedado atrapada en la glucosa y el oxígeno producidos en la reacción.

Con fines analíticos es conveniente dividir la fotosíntesis en dos fases: una reacción luminosa y otra oscura. En la fase luminosa, dos moléculas relativamente inestables captan la energía solar: el adenosín trifosfato (o sea, ATP) y el nicotinamida-adenín-dinucleótido fosfato (NADPH). En la fase oscura, el ATP y el NADPH suministran la energía necesaria para formar glucosa a partir de dióxido de carbono.

Ambas fases tienen lugar en unas estructuras de la célula vegetal denominadas cloroplastos. El cloroplasto es un orgánulo complejo que tiene su propio material genético y es capaz de sintetizar algunas de las proteínas necesarias para su funcionamiento. Su interior está ocupado por unos sacos membranosos denominados tilacoides, que en muchos cloroplastos están apilados unos sobre otros en grupos denominados grana. Es en las membranas tilacoideas donde se realiza la reacción luminosa de la fotosíntesis; la fase oscura tiene lugar en la porción soluble, no membranosa, del cloroplasto.

Dado que los enzimas que llevan a cabo la conversión del dióxido de carbono en glucosa son solubles y pueden ser estudiados fácilmente por los bioquímicos, la fase oscura fue la parte de la fotosíntesis sometida a experimentación en primer lugar. Melvin Calvin y sus colaboradores, de la Universidad de California en Berkeley, descifraron el proceso enzimático de la fase oscura de la fotosíntesis hace aproximadamente 20 años. La verdad es que dicha etapa oscura tiene lugar con mayor frecuencia en presencia de luz (cuando las hojas están iluminadas); si se denomina fase oscura es porque no requiere la participación directa de la luz. La reacción puede desarrollarse en la más absoluta oscuridad, siempre que se suministren los sustratos adecuados.

El mecanismo por el cual la fase luminosa logra la conversión inicial de la luz en energía química es el punto clave de la fotosíntesis. Sólo desde hace muy poco tiempo los investigadores de distintos campos de la biología y la bioquímica han empezado a formarse una imagen global de la organización de la reacción luminosa en la membrana fotosintética.

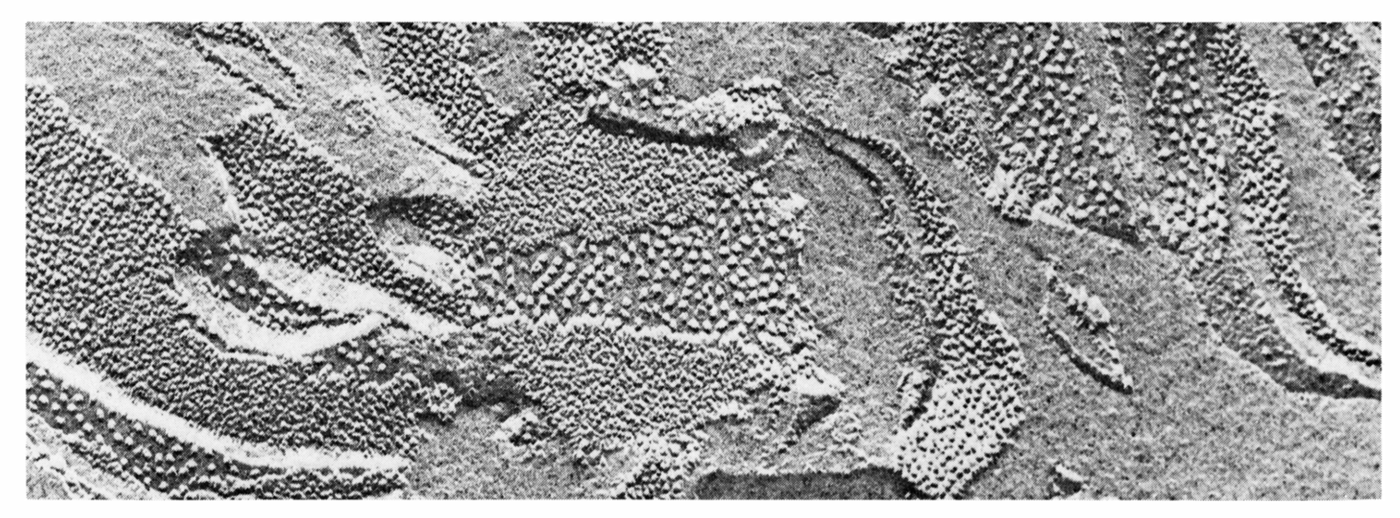
La clorofila, un pigmento verde, es la molécula central de la fotosíntesis. Ella y otros pigmentos asociados, como los betacarotenos y la xantofila, absorben la luz de gran parte del espectro visible. La absorción de un solo fotón de luz por una molécula de clorofila provoca ya una separación de cargas positivas y negativas en una región especial de la membrana denominada centro de reacción. Tras la absorción, el centro de reacción pierde un electrón, y queda cargado positivamente; el electrón libre, que es ahora el portador de gran parte de la energía del fotón absorbido, puede reducir (es decir, aumentar el número de electrones) un determinado tipo molecular de la membrana.

La capacidad de mantener separadas las cargas es una de las características básicas del sistema. Aunque son muchas las cargas positivas y negativas que hay en la membrana, existe un mecanismo que impide la reacción del electrón cargado de energía con las cargas positivas, pues esto daría lugar a la cesión de energía en forma de calor o de fluorescencia. Como veremos, la membrana tilacoidea está diseñada especialmente para evitar la pérdida de energía de los electrones excitados.

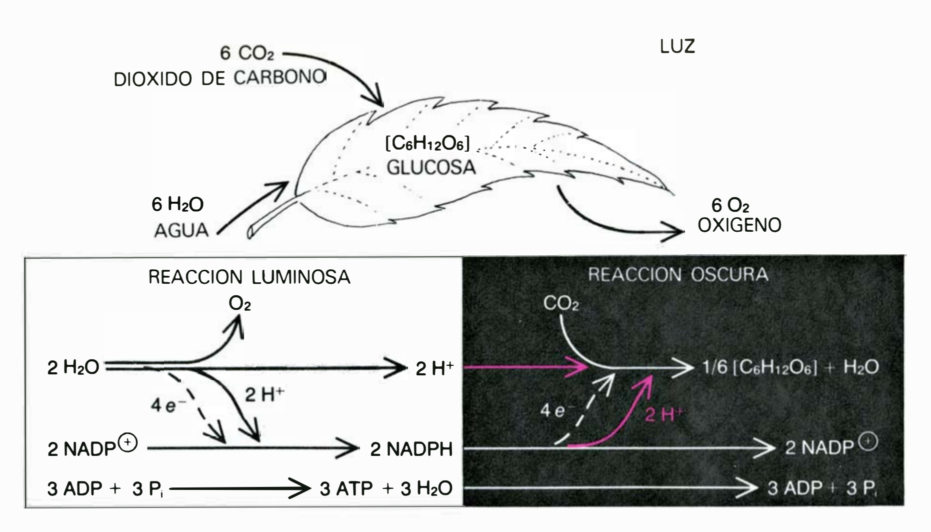
Un proceso análogo de excitación y separación de cargas se da en las células solares artificiales fabricadas a base de cristales de silicio. Mediante la adición controlada de pequeñas cantidades de impurezas se provoca en el cristal de silicio de estas células fotovoltaicas la aparición de unas zonas positivas (*p*) y otras negativas (*n*). En realidad, las zonas *n* retienen con más facilidad las cargas negativas y las *p* las positivas. Al incluir la célula fotovoltaica en un circuito, la acumulación de electrones excitados puede generar una pequeña corriente eléctrica.

Al cloroplasto le faltan los cables del circuito, pero tiene un conjunto especial de moléculas situadas en las cercanías del centro de reacción que impiden la pérdida de la energía acumulada en los electrones excitados. Estas moléculas, algunas de naturaleza proteica y otras lipídicas, dan lugar a una cadena de transporte de electrones que retira los

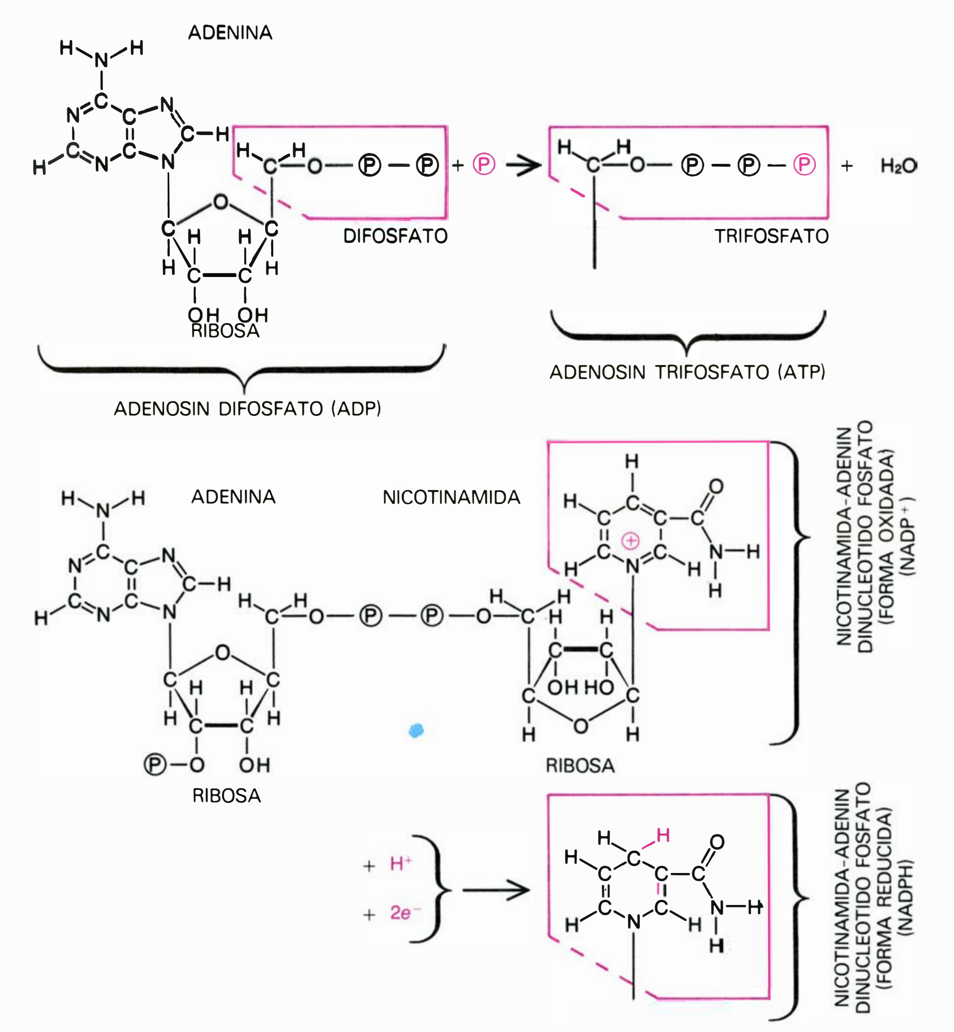
LA FOTOSINTESIS SE LOCALIZA en un orgánulo de las células vegetales denominado cloroplasto. Esta microfotografía electrónica muestra parte de un cloroplasto de una célula de hoja de maíz aumentado unas 70.000 veces. Las membranas que se aprecian muy próximas entre sí, y que en muchas células se apilan formando agrupaciones denominadas grana, llevan a cabo el paso inicial de la fotosíntesis: la "reacción luminosa", en la que se captan fotones de luz. Esta captura la realizan ciertos complejos de proteínas y pigmentos englobados en las membranas. Los complejos pueden ponerse de manifiesto utilizando la técnica de criofracturación, que provoca la división de la doble capa lipídica de la membrana en dos hojas. La separación de una de las membranas del conjunto permite contemplar partículas de gran tamaño como las que se observan en la microfotografía inferior de esta misma página. A menos que se indique lo contrario, todas las microfotografías de este artículo han sido obtenidas por el autor.



AMPLIA MUESTRA de partículas englobadas en la membrana. Quedan expuestas al someter los grana a criofracturación. La superficie fracturada es "sombreada" con platino y revestida de carbono, formándose una réplica, separable de la muestra congelada, que puede observarse al microscopio electrónico. Los complicados contornos se deben a que la fractura tiende a saltar de un nivel a otro al atravesar las membranas. Las partículas de gran tamaño situadas en el centro se visualizan al fracturarse una membrana de la zona de apilamiento. La ampliación es de 70.000 diámetros.



REACCION QUIMICA DE LA FOTOSINTESIS (*arriba*). Puede expresarse como la síntesis de glucosa (un azúcar de seis carbonos) a partir de agua y dióxido de carbono, con liberación de oxígeno. Los productos de la reacción almacenan más energía que los reactivos, de forma que la energía de la luz ha quedado almacenada químicamente. El proceso consta de dos fases distintas (*abajo*). En la primera de ellas (reacción luminosa) los fotones captados por la clorofila y otros pigmentos proporcionan la energía que descompone dos moléculas de agua en una de oxígeno, cuatro protones (núcleos de hidrógeno, simbolizados por H^+) y cuatro electrones (e^-). Simultáneamente, los cuatro electrones y dos de los protones transforman dos moléculas de nicotinamida-adenín-dinucleótido fosfato ($NADP^+$) en su forma reducida (NADPH). Otra porción de la energía lumínica sirve para la adición de fosfato inorgánico (P_i) al adenosín difosfato (ADP), con lo que se forma una molécula que almacena más energía: el adenosín trifosfato (ATP). En la reacción oscura (puede desarrollarse en ausencia de luz) la energía contenida en el NADPH y el ATP se aprovecha en la conversiónde dióxido de carbono e hidrógeno en glucosa y agua. La mitad de los protones y todos los electrones necesarios los suministra el NADPH.



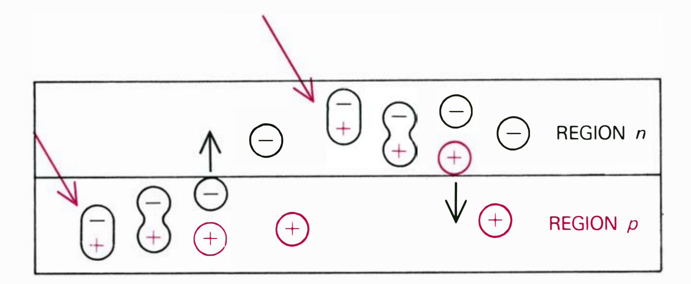
AFINIDADES ESTRUCTURALES entre los dos sistemas moleculares que sirven de almacén temporal de la energía lumínica captada por los fotopigmentos. En el sistema ADP-ATP (*figura superior*), la energía se recoge en el enlace del grupo fosfato adicional. En el sistema $NADP^+$-NADPH (*abajo*), el almacenamiento se efectúa mediante la adición de un protón y dos electrones al grupo nicotinamídico.

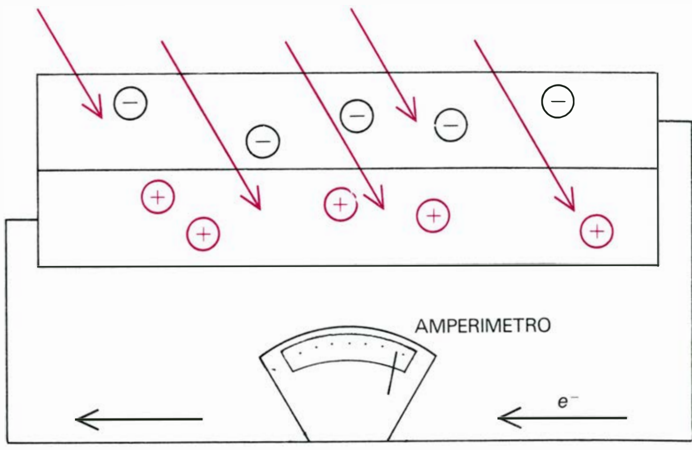
electrones del centro de reacción y, en última instancia, permite su utilización en la transformación del $NADP^+$ en NADPH. El $NADP^+$ es la forma oxidada del NADPH. Con la adición de un par de electrones excitados y de un protón, el $NADP^+$ pasa al estado reducido NADPH, una de las moléculas necesarias en la reacción oscura. Dado que los electrones no pueden fluir continuamente desde el centro de reacción hasta el $NADP^+$ sin ser repuestos, la membrana dispone también de un sistema que reduce el centro de reacción una vez que éste ha sido oxidado (es decir cuando ha perdido un electrón). El sistema extrae los electrones del agua liberando a la vez protones (hidrogeniones) y moléculas de oxígeno. Por tanto, el flujo de electrones en el tilacoide sigue un recorrido lineal desde el agua hasta el NADPH.

Una de las características de este flujo lineal es que tiene carácter vectorial, o unidireccional. Los electrones se desplazan del interior al exterior del sáculo tilacoideo. El interior del sáculo, por tanto, se carga pronto positivamente y el exterior negativamente, lo que tiene dos consecuencias importantes. Primero, el desplazamiento de electrones provoca un potencial eléctrico a través de la membrana semejante en muchos aspectos al potencial generado entre las zonas *p* y *n* de una célula fotovoltaica de silicio. Segundo, la concentración de protones es ahora sensiblemente distinta a ambos lados de la membrana, siendo el interior mucho más rico en ellos. Puesto que las fuerzas aleatorias de la difusión tienden a desplazar las moléculas de forma que se igualen las concentraciones a ambos lados de un obstáculo separador, la diferencia de concentración, o gradiente de protones, es una fuente de energía potencial. Cuanto mayor sea el gradiente protónico, mayor será la cantidad de energía potencial.

Podemos ya intuir por qué el proceso de la fotosíntesis está asociado a membranas. Es probable que la reacción luminosa tenga lugar en una membrana precisamente para mantener los centros de reacción y los sistemas de transporte de electrones en una orientación asimétrica capaz de generar un gradiente protónico. La membrana es a la vez una barrera suficientemente intensa como para mantener la diferencia de carga originada por ella misma, y por tanto, almacenar la energía durante un breve período.

Aunque en un solo cuanto de luz absorbido hay (aprovechándola al máximo) energía suficiente como para conducir cuesta arriba el flujo de electrones



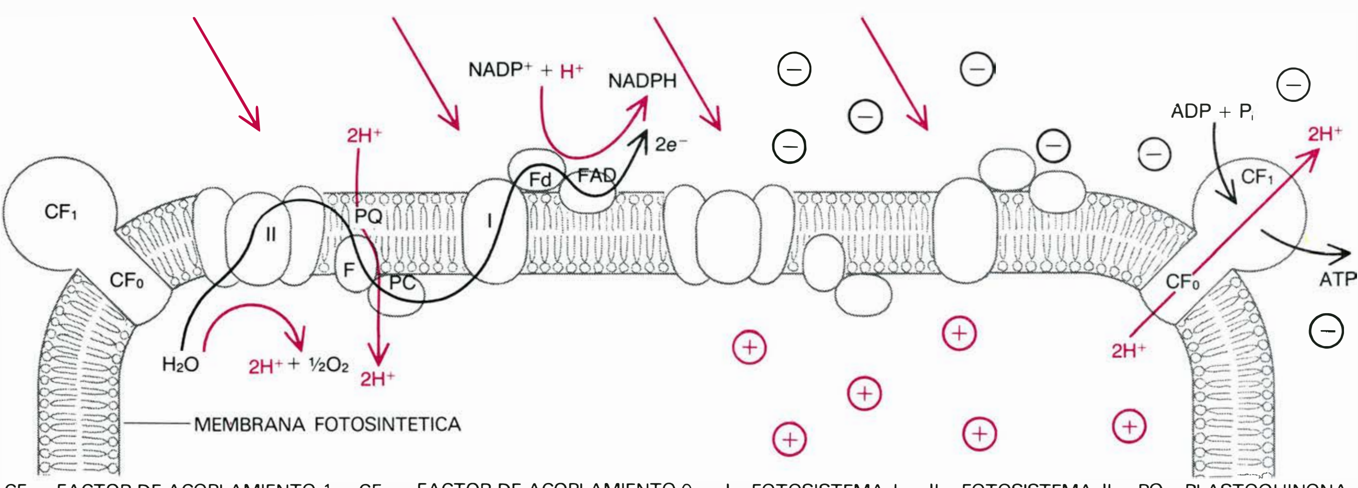


MECANISMO SEPARACION DE CARGAS que utilizan las células solares artificiales. Estas células fotovoltaicas se fabrican con cristales de silicio que dan lugar a la formación de dos capas separadas a las que se añaden trazas de otros elementos. Las "impurezas" hacen a una de las capas más afín a las cargas positivas (*p*) y a la otra más afín a las negativas (*n*). Al incidir un fotón sobre la zona *n* de la célula, una carga positiva excitada (en realidad un "hueco" o ausencia de un electrón) pasa a la zona *p*. Análogamente, la absorción de un fotón en la zona *p* provoca la translocación de una carga negativa a la región *n* (*izquierda*). Bajo iluminación continua (*derecha*), la distribución de cargas positivas y negativas a ambos lados de la unión *p* – *n* puede alcanzar el nivel de hasta 0,6 volt. La efectividad de la célula fotovoltaica depende de su asimetría y de la capacidad de la zona de unión de mantener separadas las cargas positivas y negativas. El interior del sáculo del tilacoide recuerda esta disposición y estructura.

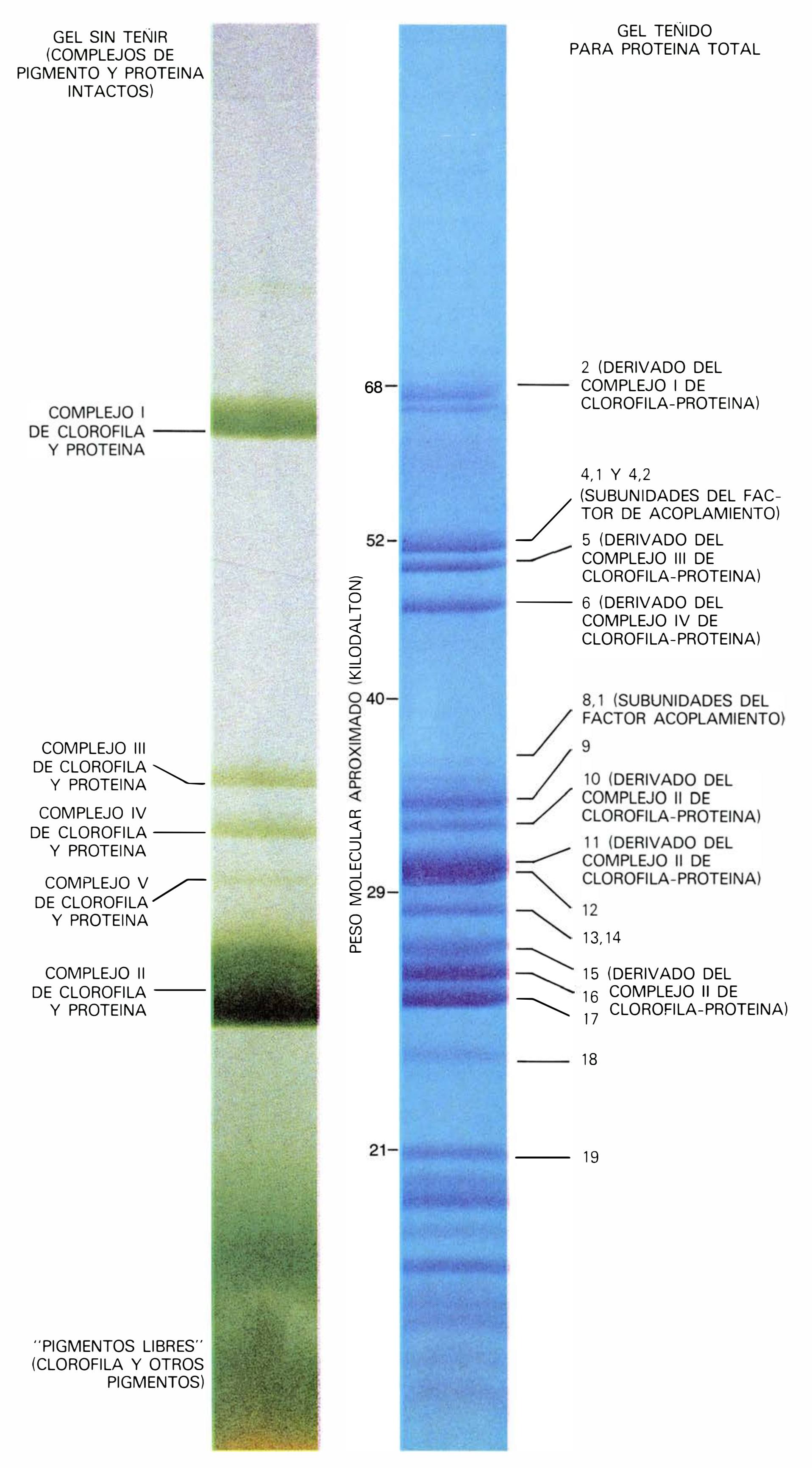
desde el agua hasta el $NADP^+$, la membrana del tilacoide necesita absorber evidentemente dos cuantos en centros de reacción distintos, unidos en serie, para llevar a término el proceso. Los dos centros de reacción, denominados generalmente fotosistema I y fotosistema II, fueron descubiertos gracias a que las moléculas de clorofila de los dos centros absorben luz de longitudes de onda ligeramente distintas. ¿Por qué emplea la planta un fotón adicional en el transporte de electrones? La respuesta quizá la hallemos en el proceso de síntesis, en el tilacoide, de la otra molécula necesaria en la reacción oscura, el ATP.

El transporte de electrones va acompañado de una acidificación del medio interno del sáculo del tilacoide. Los hidrogeniones responsables de esta acidificación parecen proceder de dos fuentes: los protones liberados durante la extracción de electrones del agua y los protones que van siendo bombeados a través de la membrana tilacoidea en el curso del transporte electrónico. Gran parte de la energía almacenada en los electrones excitados por la luz parece consumirse en esta traslocación de protones a través de la membrana, utilizándose una elevada porción de la energía restante en la síntesis de NADPH.

El bioquímico británico Peter Mitchell fue el primero en apreciar que la acumulación de protones en el sáculo tilacoideo podría servir como fuente de energía para la fosforilación (la adición de un grupo fosfato a una molécula de adenosín difosfato, ADP, para dar ATP). La fosforilación del ADP en ATP requiere energía y, por tanto, no sucede espontáneamente. Mitchell supuso que



PARA LOGRAR una separación eficaz de las cargas, la membrana fotosintética del cloroplasto se sirve de una cadena de reacciones que toman la energía de los fotones de la luz. Durante este proceso, el $NADP^+$ es reducido a NADPH. La diferencia de potencial resultante de la separación de cargas se aprovecha para llevar a cabo la conversión del ADP en ATP. Ambos procesos, que constituyen la reacción luminosa, tienen lugar simultáneamente en unos sáculos membranosos denominados tilacoides. Para mayor claridad se han representado aquí los procesos, separadamente, en un tilacoide que es parte de un granum de mayor tamaño. A la izquierda de la ilustración dos centros de reacción, denominados fotosistema I y fotosistema II, absorben luz de longitudes de onda ligeramente distintas. En el fotosistema II se disocia el agua en dos protones (H^+), oxígeno y dos electrones (e^-). Los electrones se transfieren a través de una cadena de transporte de electrones que consta de plastoquinona, plastocianina, el fotosistema I, ferredoxina y flavín-adenín dinucleótido, hasta que, con la ayuda de un protón, reducen el $NADP^+$ a NADPH en la superficie externa de la membrana del tilacoide. A la derecha observamos una acumulación de cargas positivas en la zona interna, que proceden de dos fuentes: los protones liberados por la disociación del agua y cargas positivas que parecen ser traslocadas desde el exterior a través de la membrana durante el transporte de electrones. El flujo de protones hacia el exterior a través de los factores de acoplamiento (CF_0 y CF_1) permite la síntesis de ATP a partir de ADP.



COMPONENTES DE LA MEMBRANA FOTOSINTETICA, que pueden separarse disolviendo la membrana en un detergente (dodecilsulfato de litio) y haciendo migrar los componentes proteicos en un gel de poliacrilamida en presencia de un campo eléctrico, técnica conocida como electroforesis. La velocidad de migración de una molécula depende en gran parte de su tamaño. El color verde de varias de las fracciones del cromatograma izquierdo prueba que la clorofila forma complejos con distintas proteínas de la membrana. El cromatograma derecho ha sido teñido con un colorante azul que afecta a todos los componentes polipeptídicos (proteínas y subunidades de proteínas) de la membrana. Las bandas son más numerosas y aparecen en distintas posiciones que las del gel de la izquierda, ya que los complejos de pigmento y proteína habían sido previamente disociados por efectos del calor. Ambos cromatogramas fueron obtenidos por el doctor Nam-Hai Chua, adscrito a la Universidad Rockefeller.

la generación de un gradiente electroquímico a través de una membrana mediante el transporte de protones podría servir precisamente como fuente de energía. Sugirió que ciertos enzimas especializados de la membrana tilacoidea podrían extraer la energía inherente al gradiente para llevar a cabo la síntesis de ATP. Cuando Mitchell expuso sus ideas, pocos bioquímicos estuvieron dispuestos a tomarlas en consideración. Hoy en día las líneas esenciales de la propuesta de Mitchell han sido aceptadas por la mayoría de investigadores.

Las predicciones fundamentales del modelo de Mitchell, a menudo denominado teoría quimiosmótica, han sido confirmadas experimentalmente en muchos laboratorios. Entre estos experimentos se encuentra uno de André T. Jagendorf, de la Universidad de Cornell, y Ernest G. Uribe, de la Universidad del Estado de Washington, que demuestra que el ATP puede ser sintetizado en la más absoluta oscuridad cuando se genera artificialmente un gradiente de protones a través de la membrana tilacoidea. Los trabajos realizados en el laboratorio Efraim Racker, de la Universidad de Cornell, demostraron que una proteína de gran tamaño de la superficie externa de la membrana desempeña un papel fundamental en la fosforilación. Aun cuando el transporte de electrones no se alteraba al extraer esta molécula de la membrana, el ADP no podía ya fosforilarse. Al reponer la molécula en la membrana, la fosforilación proseguía normalmente. Dado que esta molécula acopla el transporte de electrones con la fosforilación recibe generalmente la denominación de factor de acoplamiento.

Debe hacerse notar que una de las predicciones fundamentales de Mitchell fue que la membrana del tilacoide había de ser asimétrica, tanto estructural, como bioquímicamente. Una membrana perfectamente simétrica no podría generar las diferencias de concentración protónica requeridas en la fosforilación, ni tampoco podría utilizar de forma eficaz dichos gradientes. Como veremos, esta predicción ha sido demostrada en los estudios de la organización estructural de la membrana tilacoidea.

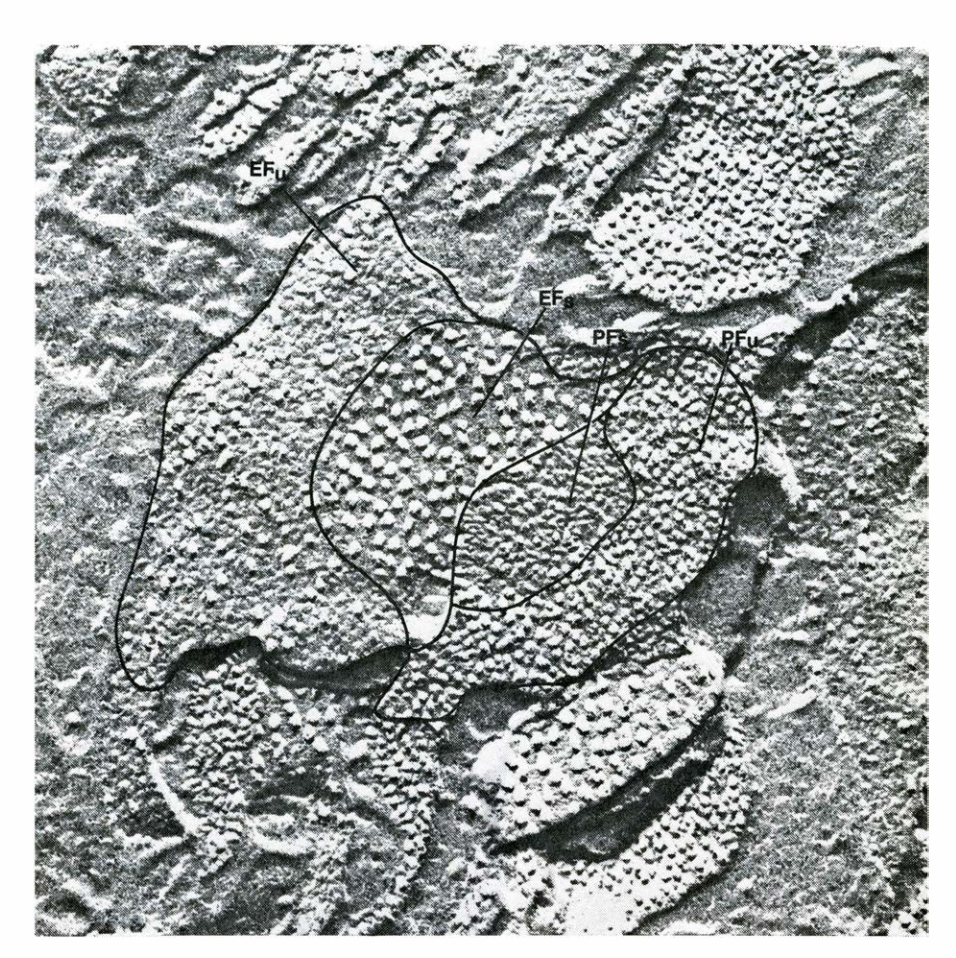
La membrana purpúrea de *Halobacterium halobium*, bacteria que habita en ambientes salinos, es la membrana transductora de energía más sencilla, y tiene una sola proteína que trasloca protones en respuesta a la luz. La membrana tilacoidea de los cloroplastos es bastante más complicada. Recurriendo al uso de electroforesis en gel, donde las

proteínas y las cadenas de aminoácidos (polipéptidos), de menor longitud que aquellas, pueden separarse en bandas de acuerdo con su peso molecular, se pueden clasificar las proteínas y los polipéptidos de la membrana tilacoidea en 40 tipos moleculares diferentes por lo menos, de los que sólo se ha logrado identificar bioquímicamente unos pocos. En las bandas identificadas se encuentran algunos de los componentes del factor de acoplamiento y también ciertos componentes de la cadena de transporte de electrones.

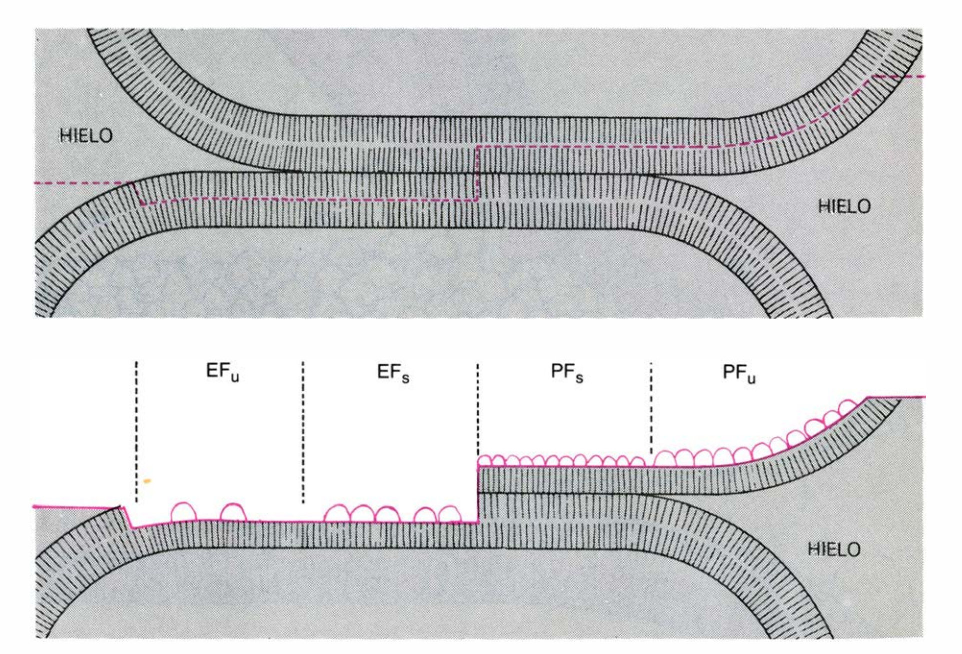
Una de las cuestiones de mayor interés para los investigadores que estudian la membrana del tilacoide es averiguar cómo se disponen en ella las moléculas de clorofila. Básicamente, la membrana consta de una doble capa de lípidos, en la que las moléculas de proteína están, bien incrustadas en ella, bien asociadas a su superficie. Las moléculas de clorofila pueden asociarse fácilmente a una doble capa lipídica, por lo que cabría considerar la posibilidad de que algunas se movieran libremente por la membrana. Investigadores de distintos laboratorios han descubierto, sin embargo, que la mayor parte del pigmento parece estar asociado a determinadas proteínas. Al disolver cuidadosamente la membrana en ciertos detergentes, varias proteínas se unen a la clorofila y aparecen como bandas verdes al someterlas a electroforesis. En estos momentos se están investigando, en diversos laboratorios y de un modo exhaustivo, las características de estos complejos de clorofila y proteínas. Se ha visto que varios de estos complejos están asociados tanto al fotosistema I como al fotosistema II.

Como mencioné anteriormente, los dos centros de reacción fotosintéticos responden a luz de longitudes de onda ligeramente distintas. ¿Implica esto que los dos centros se encuentran en centros estructurales de la membrana independientes? Aparentemente así es. Al disolver las membranas tilacoideas en un detergente suave pueden separarse de la membrana ciertos complejos de gran tamaño que contienen clorofila junto con varios polipéptidos, y que muestran actividades propias tan sólo de uno u otro de los dos fotosistemas (I o II). Este y otros experimentos similares que han sido realizados con una gran variedad de detergentes demuestran la existencia de dos centros de reacción separados físicamente.

En mi laboratorio de la Universidad de Harvard nos interesamos por el desarrollo de una descripción estructural de la membrana tilacoidea con el fin de profundizar en la comprensión de la



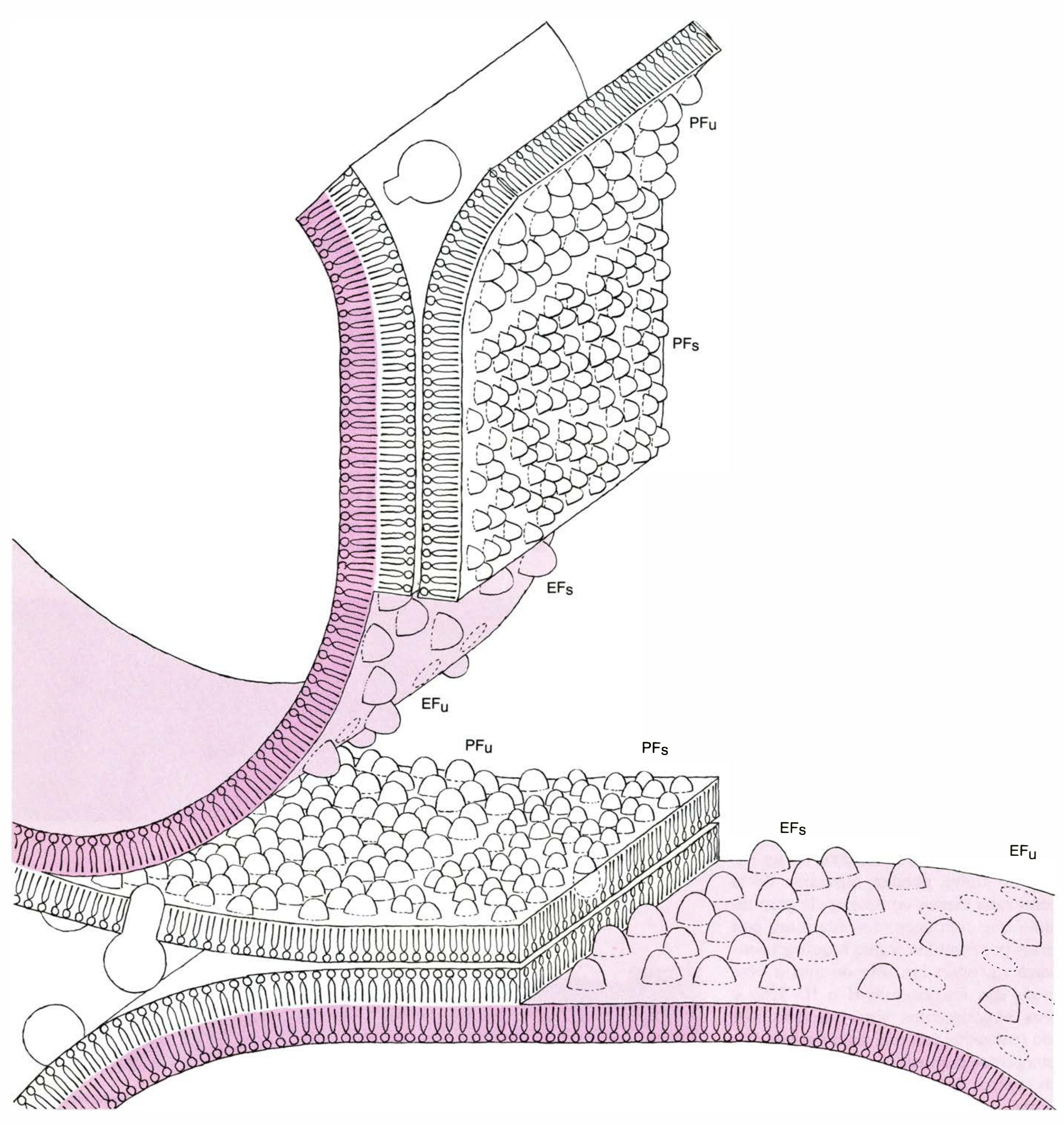
PARTICULAS DE CLASES DIVERSAS es lo que puede apreciarse en esta membrana fotosintética de cebada sometida a la técnica de criofracturación y aumentada unos 100.000 veces. El plano de fracturación saltó de la mitad de una membrana a la mitad de otra membrana adyacente a la primera y apilada a ella, lo que provoca la aparición de distintas "superficies de fracturación", denominadas EF_u, EF_s, PF_s y PF_u. EF y PF corresponden a las superficies exoplásmicas y periplásmicas respectivamente y s y u indican si la membrana estaba o no apilada. En el esquema inferior se explica su origen.



MODELO DE FRACTURACION que da lugar a los cuatro tipos de superficies de fracturación de la microfotografía superior, esquematizado. Las partículas mayores (EF_s) se encuentran en la mitad interna de una membrana tilacoidea que se halla en contacto íntimo con otra membrana. Donde las dos membranas se separan, a la izquierda, las partículas mayores son de un tamaño menor y se cuentan en un número menor; este tipo de superficie es la designada como EF_u. Donde la fractura ha dividido la membrana superior, a la derecha, las superficies que quedan expuestas son también distintas. Las superficies corresponden al interior de la mitad externa de la membrana tilacoidea y no a la mitad interna. La región PF_s es aquella en que las membranas se tocan y la PF_u, donde están separadas.

reacción luminosa de la fotosíntesis. Pese a la utilidad de las técnicas tradicionales de preparación de muestras para la microscopia electrónica, estas no nos permitieron investigar la naturaleza de las estructuras contenidas en el interior de la membrana fotosintética. Debimos acudir, por tanto, a la técnica de criofracturación, que consiste en congelar con gran rapidez una muestra, en este caso de membranas fotosintéticas, y a continuación fracturarla en una cámara de vacío. La superficie expuesta por la fractura resulta ser de gran interés para el estudio de la organización de una membrana biológica. Se hace a continuación una réplica de la superficie cubriéndola con una fina capa de platino y carbono, que se ajusta a la forma de la superficie fracturada y puede separarse luego de ésta y someterse a observación a través del microscopio electrónico. Estas réplicas reproducen con gran fidelidad los menores detalles de la muestra original. Es más, situando la fuente de platino sólo en uno de los lados de la muestra se producen unas sombras en la imagen que indican las alturas relativas de las estructuras sobresalientes, como si las observáramos en un espacio tridimensional.

El método de criofracturación permite la obtención de réplicas de dos ma-



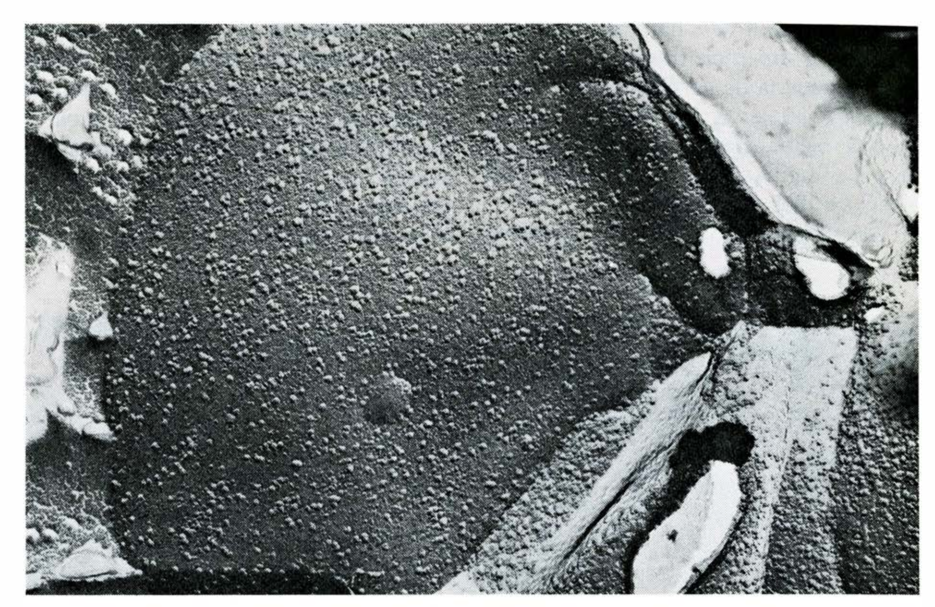
ASIMETRIA de la membrana tilacoidea, puesta de manifiesto por medio del análisis de las caras expuestas por criofracturación. Las partículas que se aprecian en la mitad interna de la membrana fracturada (*color*) tienen un aspecto distinto del que presentan las partículas expuestas en la otra mitad (*negro*). En otras palabras, la membrana muestra una polaridad distinta entre el interior y el exterior. Las partículas de gran tamaño de las caras EF_s y EF_u parecen ser estructuras compuestas por muchas subunidades asociadas al fotosistema II. Algunas partículas, PF_s y PF_u, podrían hallarse asociadas al fotosistema I. Ciertos componentes del sistema de captación de luz, porciones de la cadena transportadora de electrones y elementos de unión de la membrana localizados en el factor de acoplamiento podrían integrar parte de las pequeñas partículas de la superficie PF.

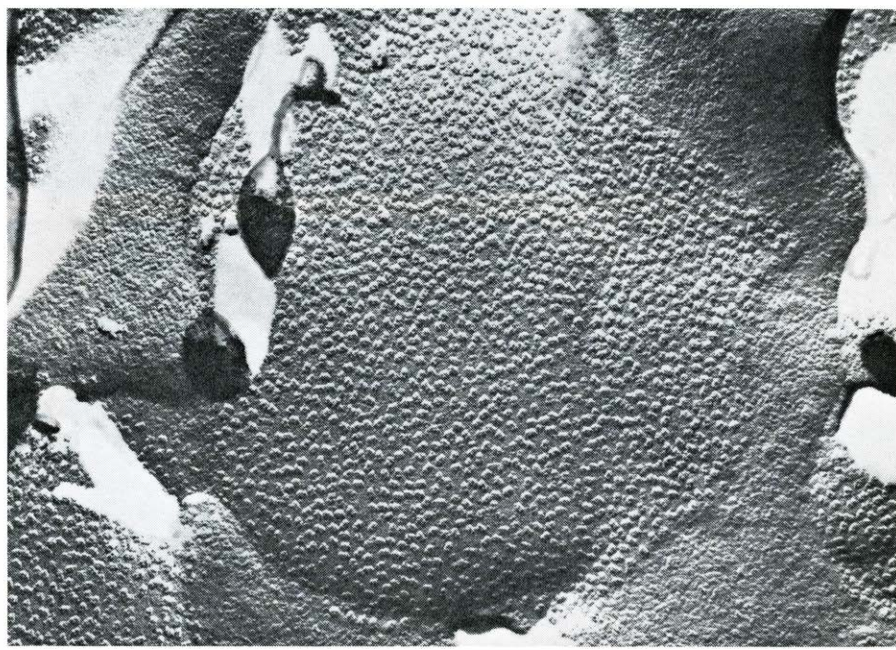
neras algo distintas. Si la muestra se somete sólo a fracturación, la doble capa de lípidos se divide por un plano central, siguiendo la fractura, separándose así la membrana en dos capas. Para determinados fines, la superficie fracturada puede grabarse ("etched") provocando la sublimación del hielo de la membrana congelada antes de obtener las réplicas. Las preparaciones grabadas son muy útiles por cuanto exponen la superficie externa de las membranas que, de otra forma, quedaría cubierta por el hielo.

A diferencia de las membranas naturales o artificiales que poseen escasa o nula asociación con proteínas, y que muestran superficies extensas y lisas, la membrana fotosintética de los tilacoides revela una sorprendente variedad de partículas de diferentes tamaños y formas. De hecho, la complejidad es tan grande, que los que nos dedicamos a su estudio empleando la técnica de criofracturación hemos tardado varios años en ponernos de acuerdo sobre qué es lo que en realidad estábamos viendo. La complejidad de la estructura da una idea de la complejidad de la función de la membrana. Se ha descubierto, por ejemplo, que la organización de la membrana es distinta en aquellas regiones donde dos membranas se encuentran muy próximas entre sí por corresponder a una zona de apilamiento. Este apilamiento, característico de la membrana tilacoidea, da lugar a las agrupaciones de grana visibles en cloroplastos intactos.

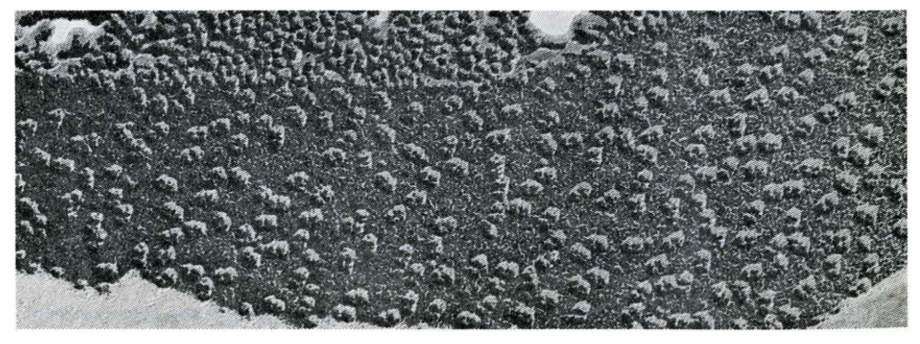
En las regiones de contacto entre membranas puede observarse una densa concentración de grandes partículas en una de las caras de fractura. En regiones adyacentes de la misma membrana, donde no hay apilamiento, la densidad del mismo tipo de partículas es mucho menor. En 1966 Seikichi Izawa y Norman E. Good, de la Universidad del Estado de Michigan, observaron que al suspender membranas tilacoideas en una solución de baja fuerza iónica las regiones de apilamiento desaparecían y la membrana se desplegaba con entera libertad. Al estudiar estas membranas por criofracturación dejan de observarse las superficies internas con gran densidad de partículas de gran tamaño.

Las membranas tilacoideas analizadas mediante la técnica de grabado dan validez a la predicción de Mitchell de que la membrana fotosintética es marcadamente asimétrica. Las superficies interna y externa de la membrana son notablemente diferentes. La superficie externa está cubierta por una miscelánea de partículas de gran tamaño (de unos 12 nanómetros de anchura) y por otras me-





SUPERFICIES EXTERNAS de la membrana tilacoidea; se logra visualizarlas a través de un proceso de congelación y grabado. En este método, las superficies de las membranas no fracturadas pero cubiertas originalmente por el hielo se hacen visibles al provocar la sublimación de éste. Las superficies expuestas se sombrean luego con platino y se obtienen réplicas por el método usual. La superficie externa de la membrana fotosintética (*arriba*) está recubierta de una miscelánea de partículas grandes y pequeñas, de 12 y 8 nanómetros de diámetro. Las partículas mayores han sido identificadas como moléculas del factor de acoplamiento. La superficie interna de la membrana (*microfotografía inferior*) está tachonada de partículas en forma de vagoneta que parecen estar constituidas por cuatro o más subunidades. El aumento con el que se han obtenido las dos microfotografías es de 55.000 diámetros.



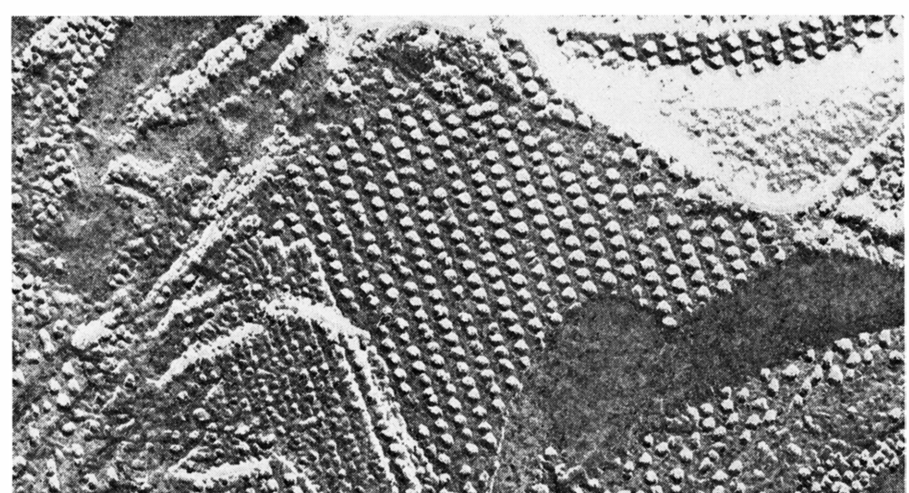
VISION AMPLIADA DE LA SUPERFICIE INTERNA de la membrana de tilacoide que muestra las partículas en forma de vagoneta con mayor detalle. El aumento es aquí de 135.000 diámetros. La partícula penetra en la membrana y puede apreciarse también en la superficie externa. Estas estructuras que atraviesan la membrana corresponden a partículas EF_s del fotosistema II ilustrado antes.

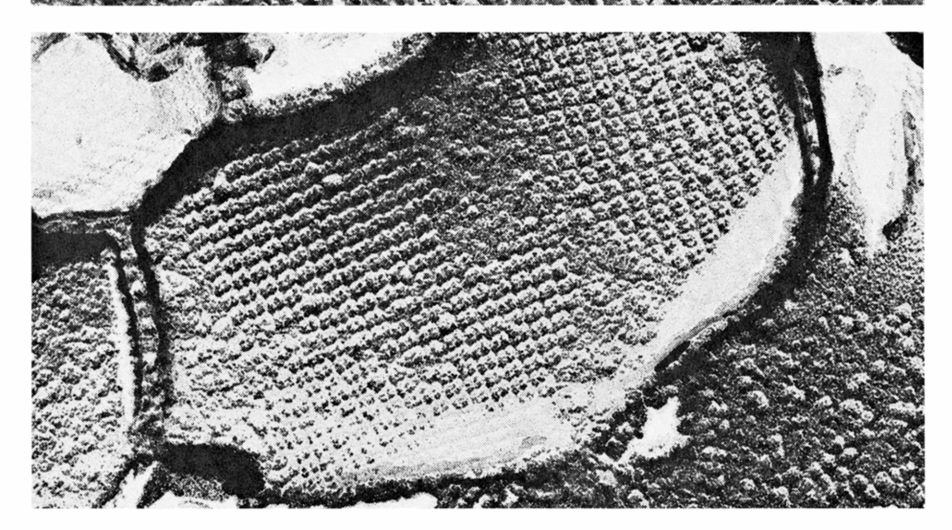
nores (alrededor de 8 nanómetros de anchura). La superficie interna está densamente cubierta de partículas cuya configuración recuerda los vagones de un tren de mercancías de unos 18 nanómetros, compuestas probablemente por cuatro o más subunidades.

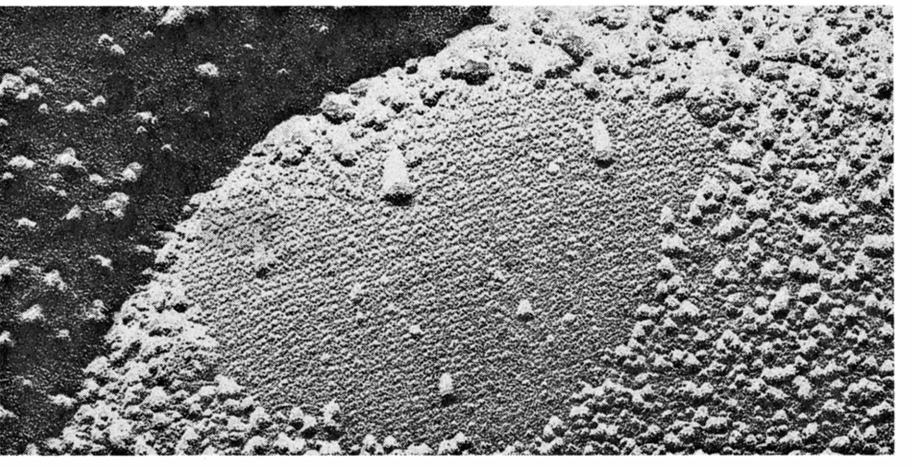
Hace algunos años, cuando trabajaba en el laboratorio L. Andrew Staehelin de la Universidad de Colorado, extraje de la membrana fotosintética el factor de acoplamiento, que sintetiza ATP a partir de ADP, a fin de apreciar el cambio de aspecto de las muestras congeladas y grabadas que provocaba su ausencia. Nuestros resultados, así como los de muchos otros laboratorios, demostraron que las macropartículas de la superficie externa de la membrana eran, en realidad, moléculas del factor de acoplamiento. Al estudiar el efecto del apilamiento y desapilamiento de la membrana sobre la densidad de estas partículas observamos, con gran sorpresa por nuestra parte, que las partículas no parecían encontrarse en la superficie externa del tilacoide, allí donde las membranas estaban próximas entre sí al formar parte de un granum; sólo se hallaban en la superficie de zonas no apiladas. Este resultado confirmaba, una vez más, la teoría quimiosmótica de Mitchell, pues demostraba que el sistema de síntesis de ATP no tenía por qué estar en conexión directa con una cadena de transporte de electrones para beneficiarse del gradiente protónico inducido por la luz.

La superficie interna de la membrana no evidencia moléculas del factor de acoplamiento; está poblada de aquella curiosa partícula de cuatro o más subunidades. Bajo ciertas condiciones, que sólo estamos empezando a comprender, estas partículas parecen "cristalizar" formando una red regular bidimensional. Tales redes no sólo se encuentran en la superficie interna de la membrana sino también en la externa, así como en preparaciones fracturadas. La partícula es visible en ambas superficies del tilacoide y también en el interior de la membrana, por lo que hay que suponer que la atraviesa. ¿Cuál es la naturaleza de esta estructura tan poco usual, y en qué consiste su asociación con el proceso de la fotosíntesis?

Las investigaciones que de forma independiente se han llevado a cabo en mi laboratorio y en el de Charles J. Arntzen, de la Universidad de Illinois en Urbana, dieron la primera prueba de que un complejo de clorofila y proteína (el CP II) estaba relacionado con la estructura que iba de una parte a otra de la membrana. Estos trabajos demostraron que la ausencia del CP II provocaba una reducción del diámetro de la partícula de gran tamaño. Ambos grupos dedujeron, en consecuencia, que esta partícula debía ser el complejo de reacción del fotosistema II. Evidentemente, el centro de reacción del fotosistema II se situaría en el núcleo de la partícula y las moléculas del CP II se dispondrían a su alrededor.

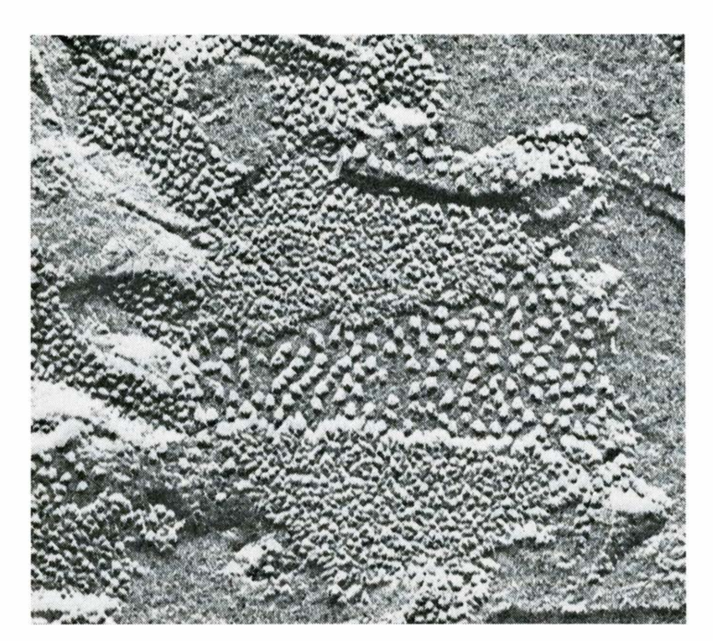


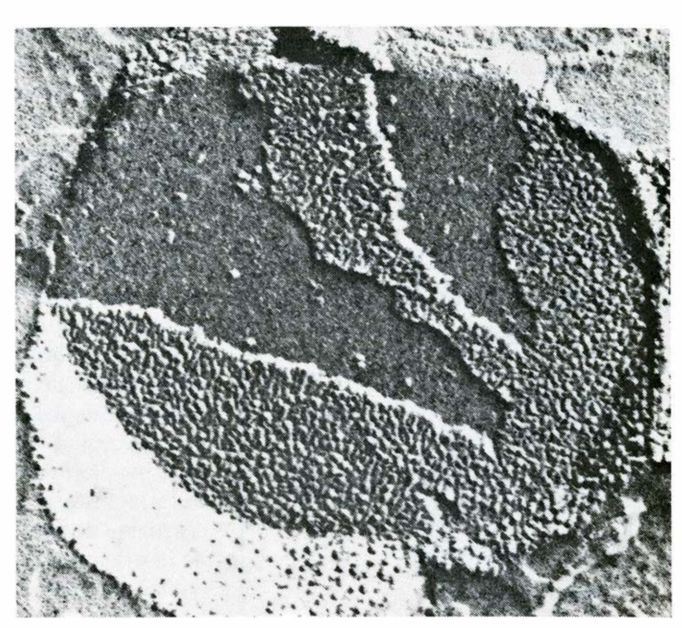




LAS PARTICULAS DAN LUGAR A REDES REGULARES en ciertas condiciones. Estas tres microfotografías, las tres con un aumento de 135.000 diámetros, representan las redes del interior de una membrana fracturada (*arriba*), de la superficie interna de una membrana intacta (*medio*) y de la superficie externa de una membrana intacta (*abajo*). En las tres, las redes tienen las mismas dimensiones. En la microfotografía inferior es menos evidente, pero se aprecia bien desde un ángulo agudo. Las redes están formadas por una partícula que atraviesa la membrana asociada al fotosistema II.

Una prueba más directa se ha obtenido recientemente del estudio de una planta de tabaco mutante, que es deficiente en la actividad del fotosistema II. Robert Cushman y yo observamos que las membranas tilacoideas del mutante carecen casi por completo de la





PARTICULAS QUE ATRAVIESAN LA MEMBRANA. Lo hacen las partículas de gran tamaño que se aprecian en la imagen izquierda de la réplica de una membrana fotosintética de una planta de tabaco normal obtenida por criofracturación. En una planta mutante que carece de fotosistema II, una réplica análoga de la membrana fotosintética no contiene las partículas de gran tamaño (*derecha*). Las restantes estructuras de la membrana de la planta de tabaco mutante parecen que no sufrieron alteración alguna. Las réplicas expuestas se hallan ampliadas 120.000 diametros.

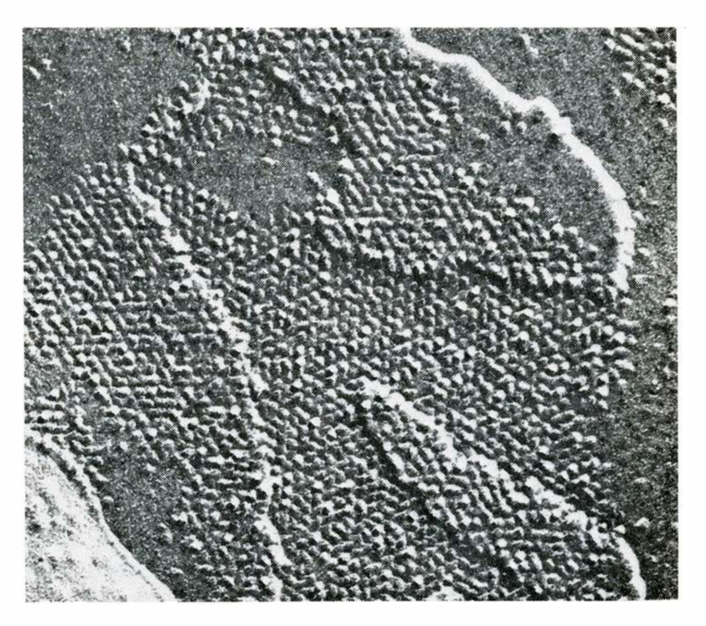
macropartícula y que el resto de las estructuras de la membrana no resultan afectadas, al menos aparentemente, por la mutación. El hecho de que estas macropartículas que atraviesan la membrana estén concentradas en las zonas de apilamiento es también coherente con la idea de que representan el complejo de reacción del fotosistema II. A través de experimentos en los que se fragmenta el sistema membranoso en fracciones de grana y estroma (o no-grana) se demuestra que la actividad del fotosistema II es mucho mayor en las membranas del grana, que contienen gran número de estas partículas.

Estudios recientes sugieren que en el apilamiento de las membranas podrían intervenir también las moléculas del CP II asociadas a esta partícula. Es evidente que cierta interacción entre el CP II y determinadas sales inorgánicas sería la responsable del proceso de apilamiento; podría ayudar, asimismo, a la disposición de las partículas en redes. Un ejemplo de ello lo tenemos en que al incluir moléculas purificadas del CP II en sistemas membranosos artificiales, las moléculas se disponen por sí mismas formando redes regulares.

¿Dónde se localiza el fotosistema I en la membrana tilacoidea? ¿Constituye también una partícula diferenciable? Y si es así, ¿cómo interactúa con el fotosistema II? Los estudios preliminares llevados a cabo en nuestro laboratorio con mutantes con deficiencias en el fotosistema I indicaron que está asociado a un subconjunto de las partículas menores



MEMBRANA ARTIFICIAL, preparada en el laboratorio del autor añadiendo complejos del fotosistema I, purificados por Michael Newman, a dobles capas obtenidas a partir de un lípido de uso comercial. Al congelar y fracturar la membrana (proceso que a lo largo del artículo hemos denominado de criofracturación) aparecieron partículas de un tamaño igual al de las que se presentan en todas las membranas naturales. Aumento x 68.000.



COMPLEJOS de clorofila y proteína (CP II) asociados al fotosistema II; forman redes regulares en esta microfotografía de una réplica obtenida por criofracturación de una membrana artificial preparada por Anette McDonnel en el laboratorio de L. Andrew Staehelin, en la Universidad de Colorado. La réplica tiene un aumento de 125.000 diámetros. El complejo de clorofila y proteína parece también promover la formación de los grana.

que se encuentran en la membrana fotosintética tras el proceso de criofracturación. Otras investigaciones corroboran ese concepto. Por ejemplo, en los sistemas híbridos que constan de una doble capa lipídica artificial a la que se ha añadido el centro de reacción del fotosistema I purificado, se observan unas partículas de un diámetro que corresponde al de las partículas observadas en la membrana natural.

Está claro que son numerosas las distintas subestructuras de la membrana tilacoidea que juegan un importante papel en la reacción luminosa de la fotosíntesis. Uno de los aspectos decepcionantes de esta historia es lo poco que puede afirmarse de momento sobre los detalles de la organización molecular de las partículas complejas puesta de manifiesto por la técnica de criofracturación. Nos gustaría conocer, por ejemplo, la disposición exacta de los componentes de la cadena de transporte de electrones, la disposición de las moléculas de clorofila alrededor de los dos centros de reacción y la disposición de determinadas proteínas que parecen desempeñar una función sobresaliente en diversas actividades de la membrana.

De momento sólo se conoce la estructura molecular de dos sistemas más relacionados con la membrana fotosintética. Uno de estos sistemas es una proteína que bombea protones en la membrana purpúrea de la bacteria *Halobacterium halobium*. Las moléculas de bacteriorrodopsina de esta membrana se disponen formando una red hexagonal que se repite regularmente. Esta característica permitió a Richard Henderson y a P.N.T. Unwin, del Medical Research Council Laboratory of Molecular Biology de Cambridge (Inglaterra), combinar la microscopía electrónica, la difracción de rayos X y la difracción de electrones para lograr cartografiar la estructura de la proteína a una resolución de 0,7 nanómetros, suficiente como para revelar detalles moleculares tales como el plegamiento de las cadenas de aminoácidos.

La segunda estructura molecular conocida con detalle actualmente es la del complejo hidrosoluble bacterioclorofila-proteína que se halla en la bacteria fotosintética *Chlorobium limicola*. Roger E. Fenna y Brian W. Matthews, de la Universidad de Oregón, han analizado cristales de este complejo de pigmento y proteína mediante difracción de rayos X. Su detallada representación de este complejo es una visión emocionante de la intrincada disposición de las moléculas de clorofila y de las cadenas de aminoácidos necesarias para la captura de la energía solar a nivel submicroscópico.

Existe, sin embargo, una razón para suponer que ninguno de estos estudios podrá aclararnos gran cosa sobre la membrana fotosintética, ya que en cada uno de estos casos se ha trabajado con un sistema muy distinto del de un verdadero sistema fotosintético. La membrana purpúrea de *Halobacterium* interviene en un bombeo de protones activado por la luz, pero lo hace sin que intervenga ningún transporte de electrones y sin liberación de oxígeno molecular. Por otro lado, aun siendo de gran interés la estructura de la proteína-clorofila de *Chlorobium*, la proteína es hidrosoluble y por tanto no se trata de una verdadera proteína membranosa como lo son las asociadas a la membrana tilacoidea. La estructura de las proteínas unidas a la clorofila de la membrana serán con toda seguridad muy diferentes.

Tanto mis colegas como yo esperamos avanzar en el conocimiento de la membrana fotosintética gracias a la aplicación de técnicas desarrolladas recientemente, en que las imágenes de difracción de sistemas moleculares ordenados se someten al análisis de Fourier. Estas técnicas parecen adecuadas para ofrecer imágenes de alta resolución sobre la disposición de los sistemas membranosos. He descrito en este artículo dos de estos sistemas ordenados, que parecen adecuados para la aplicación de esta nueva aproximación analítica: las redes ordenadas que aparecen ocasionalmente en los tilacoides de las plantas superiores y las redes regulares originadas por las moléculas de CP II purificado en las membranas lipídicas artificiales.

Hemos comenzado a investigar también ciertas estructuras regulares de la membrana fotosintética de la bacteria *Rhodopseudomonas viridis*. Por medio de estos métodos analíticos hemos obtenido imágenes de la membrana con una resolución de alrededor de tres nanómetros. El deseo de este investigador impaciente es que no transcurran muchos años hasta que estas y otras aproximaciones lleguen a revelar la estructura de la membrana fotosintética a nivel molecular. De momento uno no puede menos que maravillarse ante la intrincada destreza del artífice que ha dado forma a esta membrana y la ha equipado para operar con una eficiencia tan exquisita. Por muy interesantes y variados que hayan llegado a ser otros sistemas vivientes, siempre vale la pena recordar que la membrana fotosintética ha permitido la aparición de tan extraordinaria diversidad biológica.

Los lagos de lava del Kilauea

Las erupciones de ese volcán hawaiano dejan charcos de basalto fundido que pueden tardar hasta 25 años en solidificarse. Proporcionan, pues, un laboratorio natural para el estudio de la naturaleza del magma del manto

Dallas L. Peck, Thomas L. Wright y Robert W. Decker

El magma –roca fundida– del interior de la tierra es el responsable de muchos fenómenos que ocurren en la superficie terrestre. El flujo magmático que asciende por entre los valles de fractura mesoceánica se va agregando a las placas rígidas que constituyen la superficie terrestre, placas que llevan a sus espaldas los continentes. Al agregarse, el magma presiona y aumenta más la distancia entre los antiguos bordes de placas. Por definición, todas las rocas ígneas se forman por enfriamiento de magma. Si éste hace erupción en superficie en forma de lava, origina rocas ígneas extrusivas, como el basalto; pero si cristaliza lentamente bajo la superficie da lugar a rocas ígneas intrusivas, como el granito. A pesar de la importancia del magma, es mucho lo que todavía ignoramos sobre él. La mayoría de los estudios en torno al enfriamiento, cristalización y otras propiedades suyas se han centrado en el análisis en laboratorio de muestras pequeñas y en extrapolaciones teóricas a partir de lava ya solidificada. Pero un grupo de investigadores, pertenecientes al Servicio Geológico de los Estados Unidos, a los Laboratorios Sandia y a diferentes universidades, hemos optado por un método distinto. Hemos estudiado lava fundida y en proceso de solidificación *in situ* examinando tres "lagos" de lava formados a consecuencia de las erupciones del volcán Kilauea, en Hawai.

Los lagos son de lava basáltica. El basalto es la roca más frecuente entre las originadas por solidificación del magma, arrojado por extrusión a la superficie de nuestro planeta, de la Luna y quizá también de otros cuerpos del sistema solar. La lava basáltica, que es rica en calcio, magnesio y hierro, cristaliza formando rocas oscuras, constituidas principalmente por feldespato plagioclasa, piroxeno y olivino (cuya gema es el peridoto), que son óxidos silíceos, por la magnetita, que es óxido de hierro, y por la ilmenita, que es óxido de hierro-titanio. Hay basalto en todos los continentes y cubre vastas extensiones como las mesetas de los ríos Columbia y Snake, en el noroeste de los Estados Unidos, y la meseta de Deccan, en la India occidental. Las lavas basálticas, que hacen erupción en los valles de fractura mesoceánicos y crean el fondo subyacente a los sedimentos de las cuencas oceánicas, han estado manando a lo largo de la historia de la Tierra, desde principios del Precámbrico hasta nuestros días. Aunque los basaltos presentan variaciones significativas en su composición química y mineral, se han formado todos a temperaturas elevadas. Por lo que, al abordar el estudio de los lagos de lava, nos propusimos analizar, entre otras cosas, la posibilidad de obtener energía geotérmica, directamente, a partir de masas profundas de magma basáltico. Si bien, en principio, la alta temperatura de la roca fundida hace atractivo su aprovechamiento como fuente de energía, en la práctica son muchos los obstáculos que dificultan la explotación de su calor.

La cifra de volcanes activos de la Tierra rebasa el medio millar. La mayoría de los cuales son predominante o completamente basálticos, entre los que se cuentan los volcanes activos que constituyen los dos tercios meridionales de la isla de Hawai: el Mauna Loa, el Hualalai y el Kilauea. El Mauna Loa, que es el mayor de los tres, se eleva cuatro kilómetros sobre el nivel del mar y casi nueve kilómetros sobre el fondo oceánico adyacente. El Kilauea es el menor, con una altura de sólo 1,3 kilómetros sobre el nivel del mar; su forma recuerda la de un plato invertido, y está perforado en su cima por una caldera, o gran cráter, de la que irradian dos zonas de fractura. Las erupciones del Kilauea suelen limitarse a la caldera y a las zonas de fractura, en particular a la zona oriental y al Halemaumau, "la hoya de fuego" de la caldera. Desde 1952 el Kilauea ha venido haciendo erupción, por término medio, una vez al año como mínimo.

El magma basáltico que alimenta las erupciones procede del manto terrestre, desde una profundidad de unos 50 kilómetros, por lo menos, bajo la superficie terrestre. De acuerdo con los datos geológicos y geofísicos, el magma que asciende desde esas profundidades se halla almacenado en un depósito, de configuración irregular, situado a unos cuatro kilómetros de la cumbre del Kilauea. En el proceso de formación de un lago de lava, ésta emerge desde el depósito, irrumpe en la superficie y fluye hacia una depresión. A veces, los diques naturales que canalizan el flujo de lava hacia el lago se derrumban, el lago queda desconectado de la fuente de lava y ésta comienza a solidificarse.

Tres fueron los lagos de lava del Kilauea que investigamos; por orden de formación, el primero surgió en 1959 en Kilauea Ibi Crater, el segundo, en 1963, en Alae Crater y, el tercero, que se constituyó en 1965, en Makaopuhi Crater. Los lagos tienen una profundidad considerable, razón por la cual tardan mucho tiempo en enfriarse y solidificarse. El período de enfriamiento varía entre diez meses y medio para el lago más somero (Alae Crater) y unos 25 años para el lago más hondo (Kilauea Iki Crater). Las diferencias entre los lagos no se reducen sólo a la profundidad, sino que se extienden también a la forma en que se llenaron, lo que incide en su temperatura, homogeneidad y estructura.

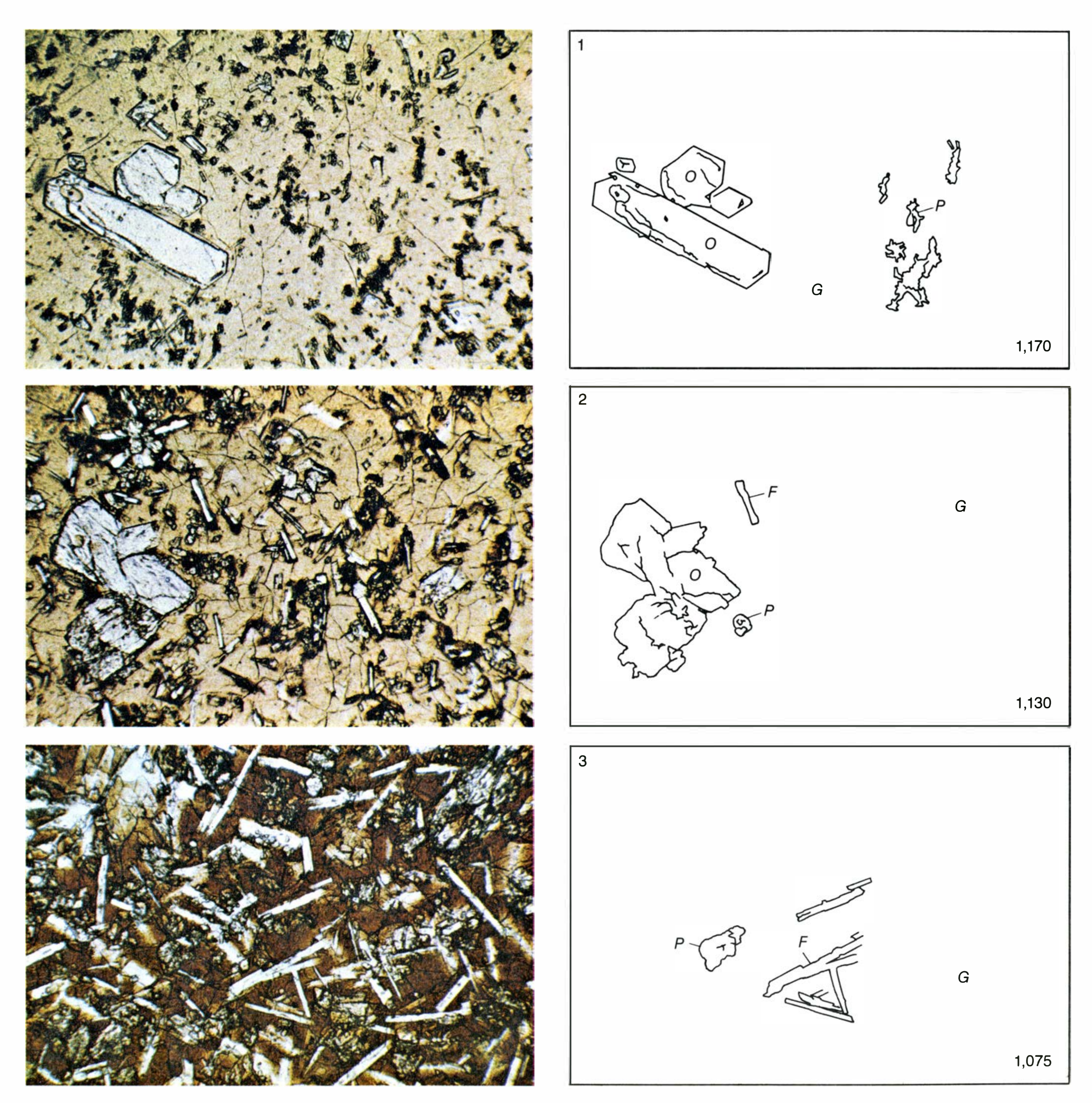
El lago de lava Kilauea Iki, que alcanza profundidades de hasta 120 metros, se formó en el transcurso de un episodio espectacular, caracterizado por 17 erupciones distintas a lo largo de 36 días. Entre erupción y erupción, la actividad volcánica cesaba, y parte de la lava se escapó del lago por una grieta abierta hacia la mitad de la ladera de depresión. Cayeron entonces fragmentos de corteza que se habían formado con anterioridad, todo lo cual determinó la aparición de una notable heterogeneidad en los niveles superiores del lago. Las temperaturas de erupción, comprendidas entre 1200 grados Celsius y 1215, fueron pues muy elevadas, y el abun-

HUNDIMIENTO DE LAVA solidificada en la superficie del lago Makaopuhi del Kilauea. Los gases procedentes de la lava fundida subyacente ascendieron hacia la superficie y quedaron atrapados bajo la corteza del lago, solidificada aunque caliente. En la fotografía superior, el gas aprisionado dio lugar a una situación gravitacionalmente inestable, que provocó el hundimiento de un fragmento de corteza cerca de la parte central de la superficie del lago. La lava que se hundió fue reemplazada por lava fundida incandescente (*centro*). En la fotografía inferior, tomada 10 minutos después, el hundimiento se había extendido a todos los márgenes de la parte activa del lago, si exceptuamos la zona del extremo superior, a la derecha. La lava adquiere un color amarillo cuando su temperatura sobrepasa los 1200 grados Celsius; es roja a temperaturas de 900 a 1100 grados y rojo oscura o negra, a menos de 900 grados Celsius. Ambas fotografías se tomaron en 1965. Las rocas ígneas extrusivas se originan por erupción del magma.

dante óxido de magnesio favoreció la formación de olivino.

El lago de lava Alae, con una profundidad de 15 metros, es pobre en olivino. Su lava se formó en el transcurso de una erupción de tres días a temperaturas comprendidas entre 1140 y 1160 grados Celsius. El lago de lava Makaopuhi tiene valores de profundidad, temperatura, formación y contenido en olivino, intermedios entre los números mencionados para los otros dos lagos. En efecto, mide una profundidad máxima de 83 metros, y se formó a raíz de una erupción de 10 días de duración a una temperatura entre 1150 y 1200 grados C. Nuestra investigación de los lagos Alae y Makaopuhi quedó bruscamente interrumpida en 1969 por nuevas erupciones que inundaron los embalses con lava nueva. El lago Kilauea Iki no ha sufrido ninguna alteración y la investigación concerniente al mismo sigue su curso.

Una vez desconectados de su fuente de lava, los lagos se fueron enfriando, proceso que nosotros examinamos efectuando más de 50 perforaciones entre 1960 y 1977 y tomando muestras testigo de material sólido y fundido. Al principio trabajábamos con un pequeño taladro rotatorio de mano accionado por un motor de sierra de cadena. Pero más tarde utilizamos ya un gran equipo de perforación montado sobre un remolque. La perforación reveló la existencia de una interesante estructura en la zona

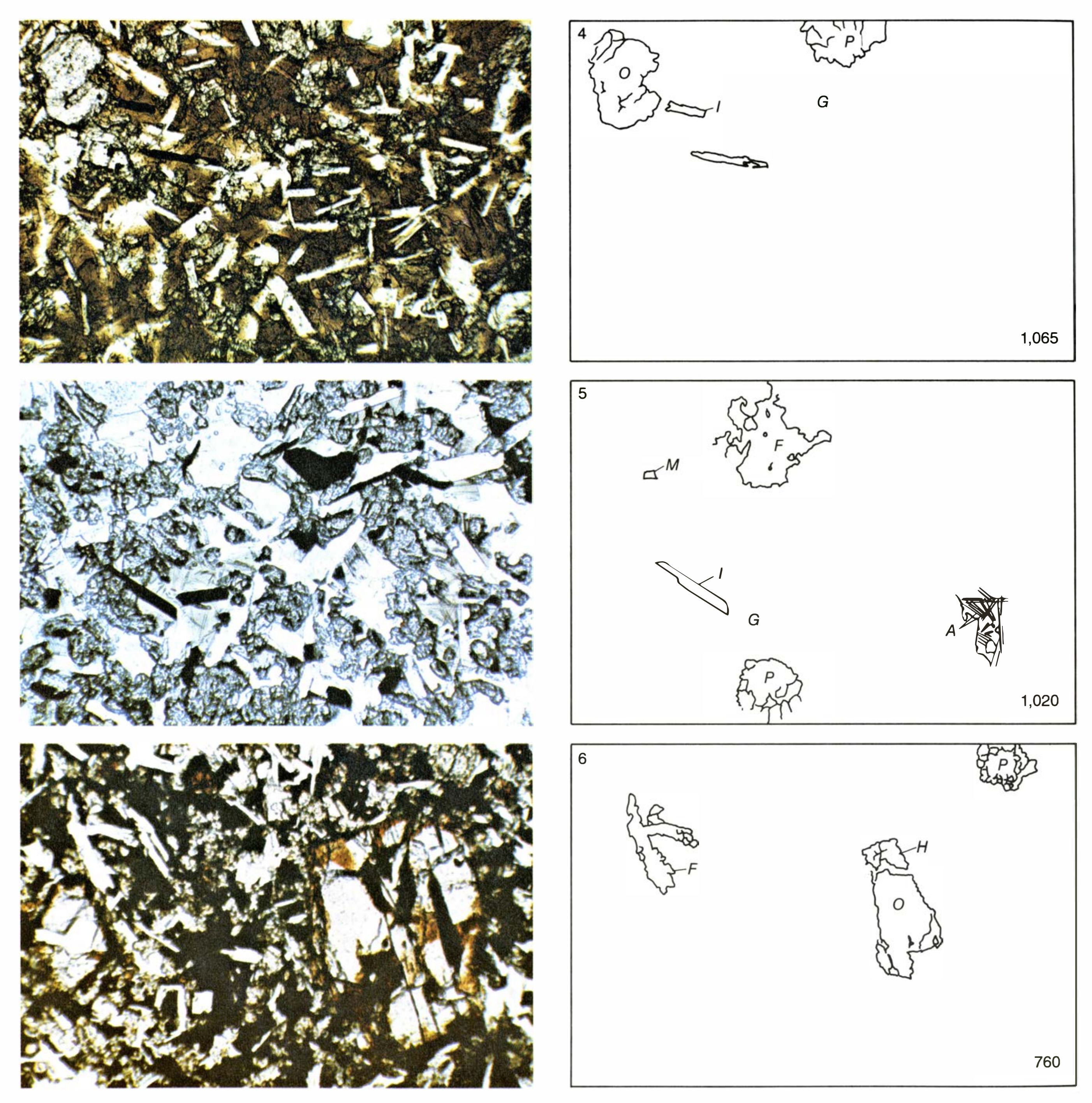


MICROFOTOGRAFIAS DE LAVA BASALTICA del lago Makaopuhi, enfriada a partir de diferentes temperaturas. Nos ponen de manifiesto cómo cristaliza la lava. A la derecha de cada fotografía hay una representación gráfica que permite la identificación de los minerales de la lava; también indica la temperatura de la lava en grados Celsius. Entre los minerales se encuentran los óxidos de silicio feldespato plagioclasa (*F*), piroxeno (*P*) y olivino (*O*), el óxido de hierro magnetita (*M*), el óxido de hierro y titanio ilmenita (*I*) y el fluofosfato cálcico apatito (*A*). Las representaciones indican también la posición de la lava vítrea (*V*), líquido superenfriado en el cual las moléculas de óxido de silicio no se han organizado en cristales. En las microfotografías *1* y *2*, que muestran lava fluida a una profundidad intermedia del lago, los cristales escasean y, entre ellos, predominan los cristales grandes de olivino (*O*). A medida que baja la temperatura por la disminución de la profundidad, aumenta la abundancia de los cristales finos

profunda. Lo mismo ocurre con la congelación de un lago de agua, el lago de lava se solidifica de arriba abajo. La superficie de un lago de lava se enfría por acción del aire, por la lluvia en particular, que en Hawai es copiosa. (En un período de cuatro años, la precipitación sobre el lago de Alae fue de 10 metros.)

Mas, a diferencia de lo que se observa en el lago de agua, el de lava se solidifica también de abajo arriba. Ello se debe a que la roca subyacente a la lava del lago está más fría que ésta. Por consiguiente, la lava fundida queda aprisionada entre dos capas de corteza solidificada y adquiere una forma lenticular.

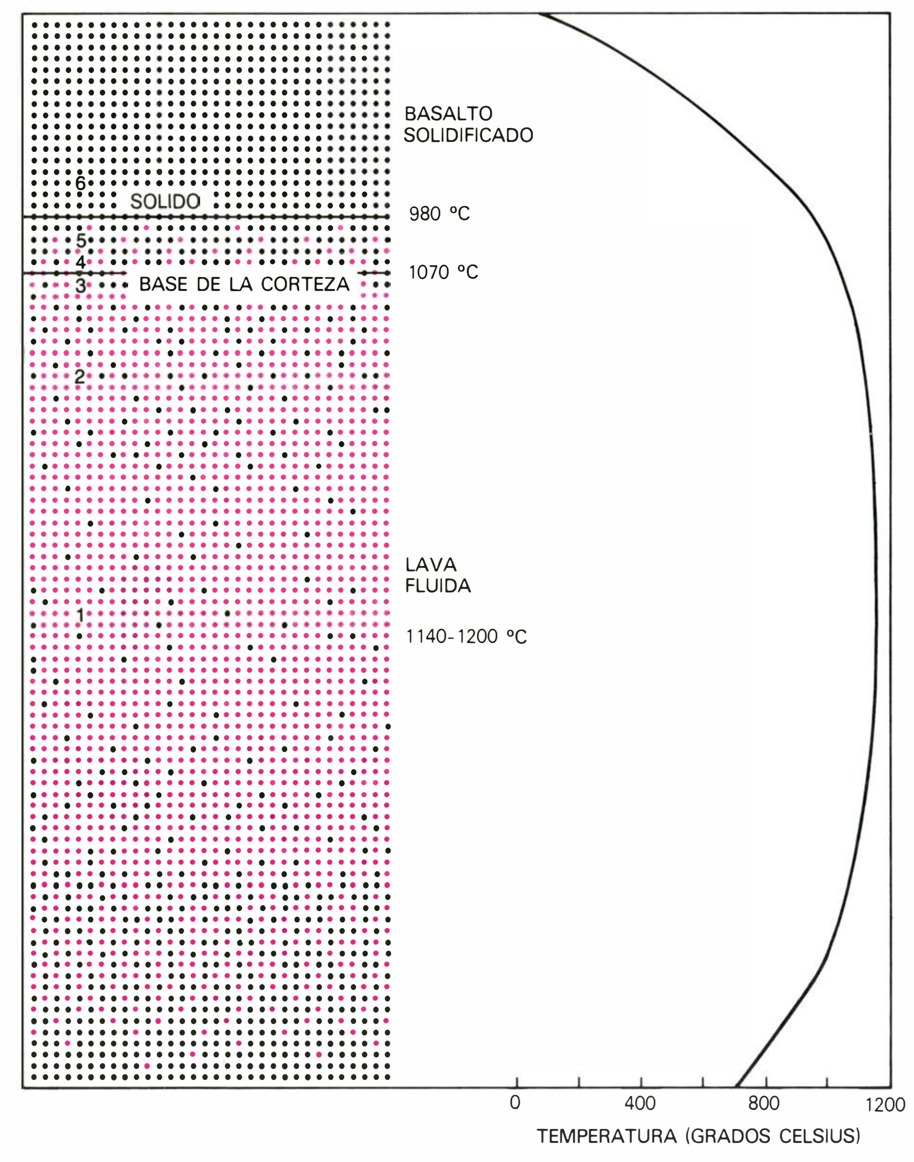
La corteza del fondo del lago es siempre más delgada que la corteza de la superficie, porque la roca subyacente tiene una conductividad térmica baja y no se halla sometida al rápido enfriamiento que provoca el agua de lluvia. A medida que el lago se va solidificando de arriba abajo y de abajo arriba, disminuye el espesor de la lente de material fundido. Al igual que los lagos de agua, los de lava son más someros en los bordes que en el centro, de suerte que la corteza profunda y la superficial se unen en los bordes del lago, aislando la lente fundida con respecto a la roca que limita la cuenca lacustre. Conforme aumenta el grosor de ambas cortezas, la lente reduce su diámetro y su espesor, acaba por desaparecer y el lago se convierte en una masa única de lava solidificada.



de piroxeno (*P*) y de los cristales pálidos y de aspecto bandeado de feldespato plagioclasa (*F*). En las microfotografías *3* y *4*, hechas en lava solidificada, extraída de la base de la corteza superior del lago, los cristales de piroxeno (*P*) y de feldespato plagioclasa (*F*) son de mayor tamaño, al par que más abundantes, constituyendo un 50 por ciento de la lava. En la *4* pueden verse los primeros cristales de ilmenita (*I*) en forma de láminas opacas. En la microfotografía *5* casi toda la lava es cristalina, constituida por un mosaico de cristales de piroxeno (*P*), feldespato plagioclasa (*F*) e ilmenita (*I*), cementado por una fina red de escaso vidrio pardo claro (*V*) que contiene agujas de apatito (*A*) y unos cuantos cubos minúsculos de magnetita (*M*). En la *6* toda la lava es sólida, pero no completamente cristalina; contiene algunas bolsas de vidrio rico en sílice. Las lavas de los lagos Alae y Kilauea Iki tienen una composición similar a la del lago Makaopuhi, y entre ellas hay ligeras diferencias químicas poco significativas.

El hielo que cubre un lago de agua está netamente diferenciado del agua subyacente. No es éste el caso de los lagos de lava basáltica, pues la lava no está constituida por un único compuesto químico, sino por distintos minerales que cristalizan a temperaturas y velocidades diferentes. La corteza superficial solidificada pasa gradual y lentamente a lava fluida a través de una región de corteza parcialmente fundida, cuyo espesor es de varios decímetros. En esta región, la temperatura y la proporción entre granos cristalizados y masa fundida aumentan suavemente con la profundidad. Con todo, hay un nivel de separación dentro de la lava parcialmente fundida que señala un cambio brusco de las propiedades físicas de ésta. Por encima de él, la lava posee la rigidez suficiente como para ser perforada en un sondeo. Por debajo, la lava es un fluido que se comporta como una pasta cuando se introduce en ella un tomatestigos. En esta interfase, cuya temperatura es de 1070 grados C., existe la misma proporción de granos cristalinos que de masa fundida. A 980 grados la lava está totalmente cristalizada, a excepción de una pequeña fracción de material en estado vítreo, correspondiente a un líquido superenfriado, en la cual las moléculas de óxido silícico no se han organizado en cristales.

La velocidad de engrosamiento de ambas cortezas (superficial y profunda) disminuye con el tiempo, porque el material solidificado es un mal conductor del calor y actúa de aislante. La corteza superior alcanzó en los tres lagos un espesor de 0,3 metros en un día, y de dos metros en un mes. Al cabo de un año la corteza superficial tenía ocho metros de espesor, en los dos lagos más produndos. En 1976, es decir, diecisiete años después de la formación del lago de lava Kilauea Iki, la corteza superior había alcanzado un grosor de 45 metros. En base a los datos de nuestros testigos, encontramos que la velocidad de aumento de espesor es proporcional a la raíz cuadrada del tiempo, hasta que la temperatura máxima del lago desciende notablemente por debajo de la temperatura inicial del centro del lago. El espesor en metros es aproximadamente igual a 0,4 multiplicado por la raíz cuadrada del tiempo expresado en días.

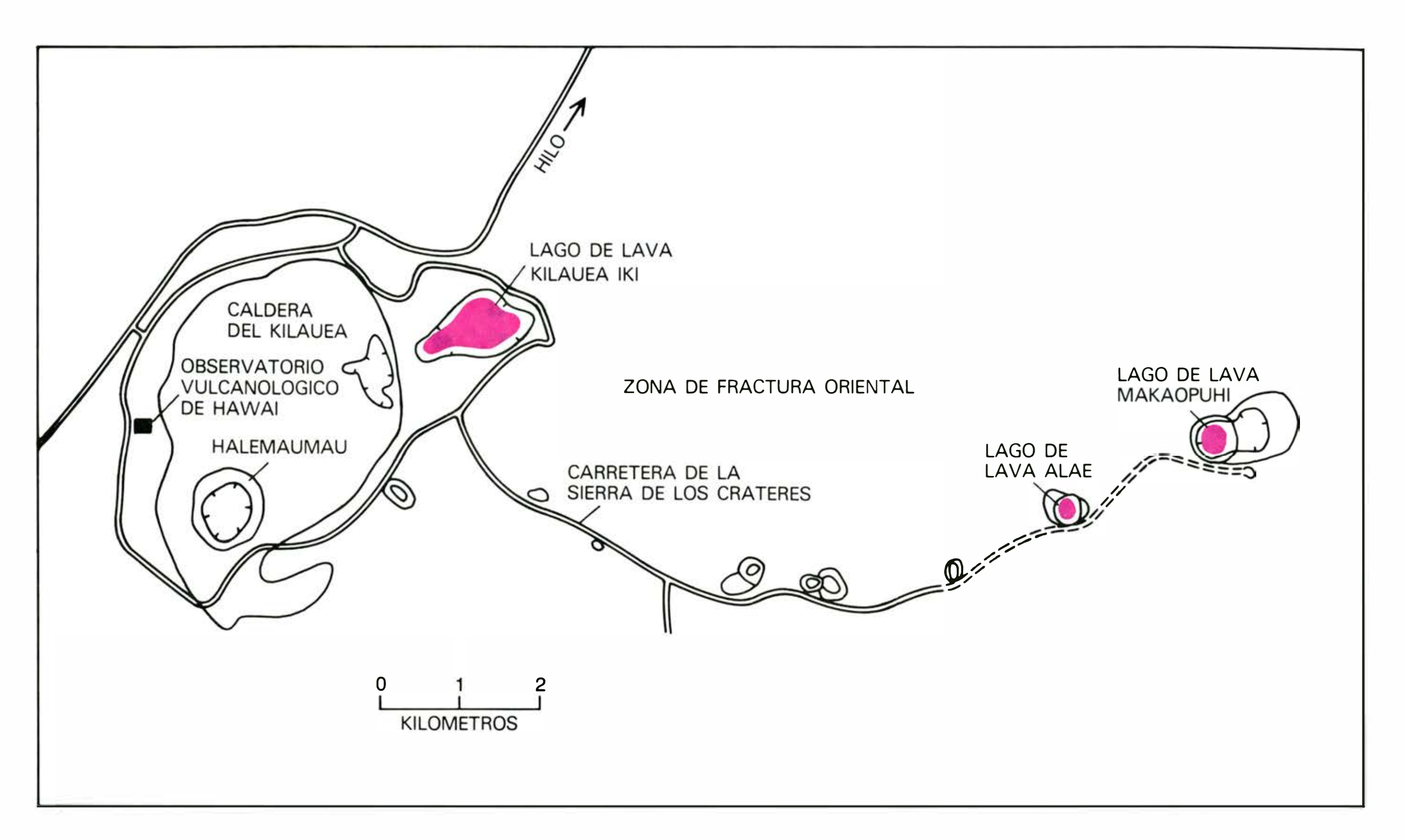


PROPORCION RELATIVA entre cristales y masa fundida en un lago de lava, representada en un corte esquemático (*izquierda*). Un lago de lava se solidifica de arriba abajo y viceversa. A temperaturas inferiores a los 980 grados C., el lago se halla completamente solidificado. Entre 980 y 1070 grados, la lava está fundida en parte, pero es rígida porque la masa fundida queda aprisionada en los intersticios de un mosaico cristalino de lava solidificada. A 1070 grados hay la misma proporción de lava fundida (*puntos de color*) que de granos cristalinos (*puntos negros*). Por encima de los 1070 grados la lava es un fluido que, cuando se la perfora con un tomatestigos, se comporta como gelatina de caramelo. Los números corresponden a las posiciones de la lava ilustrada en las microfotografías de las dos páginas precedentes. El gráfico de la derecha representa la variación de la temperatura con la profundidad.

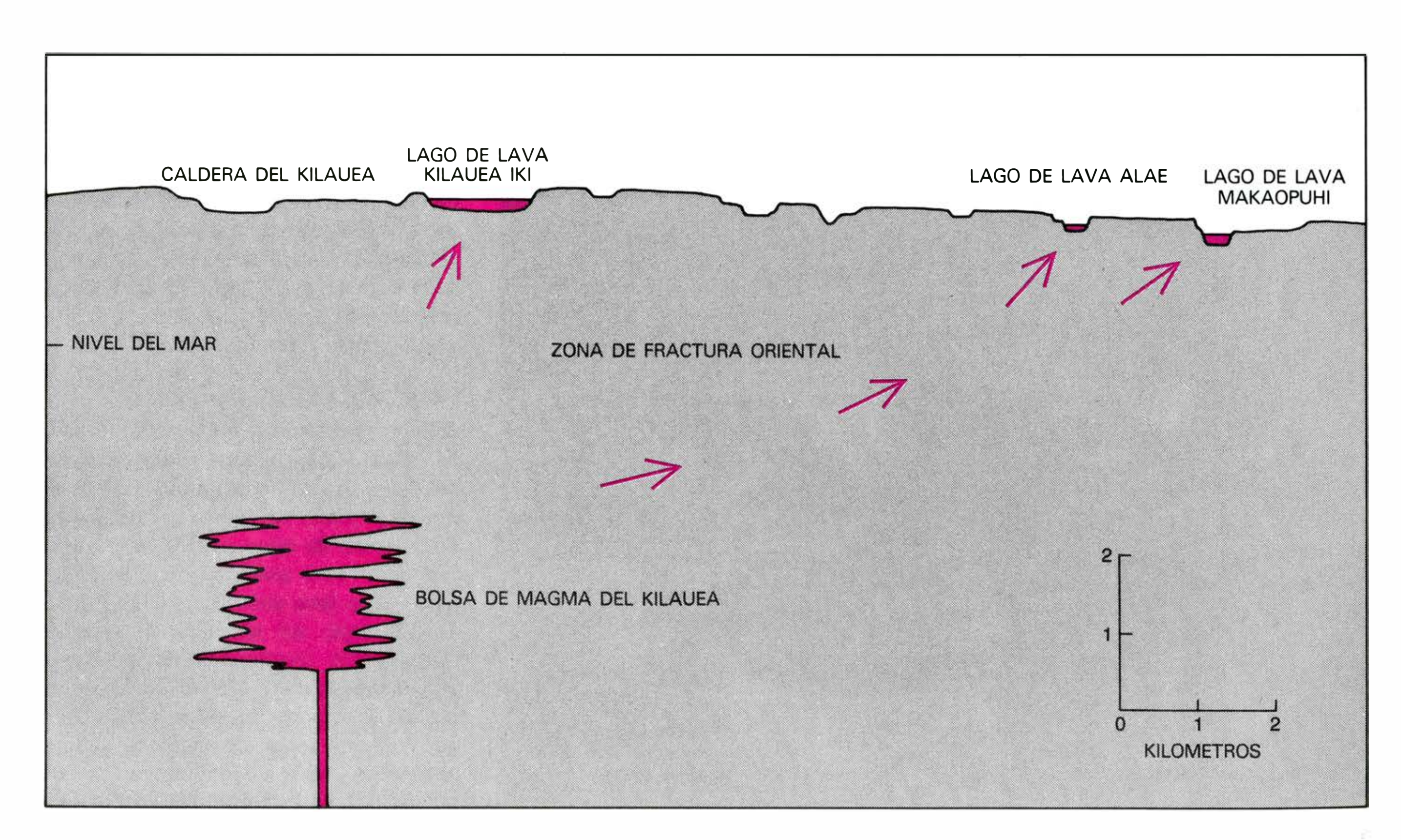
Puesto que con el transcurso del tiempo la corteza superior se engrosa más lentamente, su temperatura también debe disminuir a un ritmo menor. Las mediciones efectuadas con termopares revestidos de acero, introducidos en los pozos de sondeo, demostraron en efecto que el ritmo de disminución de la temperatura a cualquier profundidad dada de la corteza superior decrece con el tiempo. Con la ayuda de un ordenador desarrollamos un modelo numérico de predicción de la temperatura para cualquier nivel de un lago de lava. El modelo se funda en la ley de Fourier de la conducción del calor, que dice: el flujo de calor entre dos puntos es proporcional a la diferencia de temperatura entre ambos. El modelo incorpora los procesos externos que afectan al flujo de calor en un lago de lava (enfriamiento por agua de lluvia), el calor latente del basalto (calor liberado por cristalización) y las variaciones de densidad, porosidad y conductividad térmica de la lava.

Aplicamos este modelo numérico a la corteza superior del lago de lava de Alae, donde habíamos ido midiendo la temperatura de manera sistemática a lo largo de cuatro años, durante los cuales el lago se solidificó y se enfrió hasta alcanzar temperaturas inferiores a los 100 grados C. La diferencia media entre la temperatura estimada y la medida no llegó a dos grados. Esta concordancia tan notable tiende a confirmar la validez del modelo, cuyo rasgo más significativo es el de que la transformación en vapor del agua de lluvia que cae sobre la corteza caliente de un lago incide notablemente en la disminución de temperatura del lago.

Conforme se va enfriando la lava, se



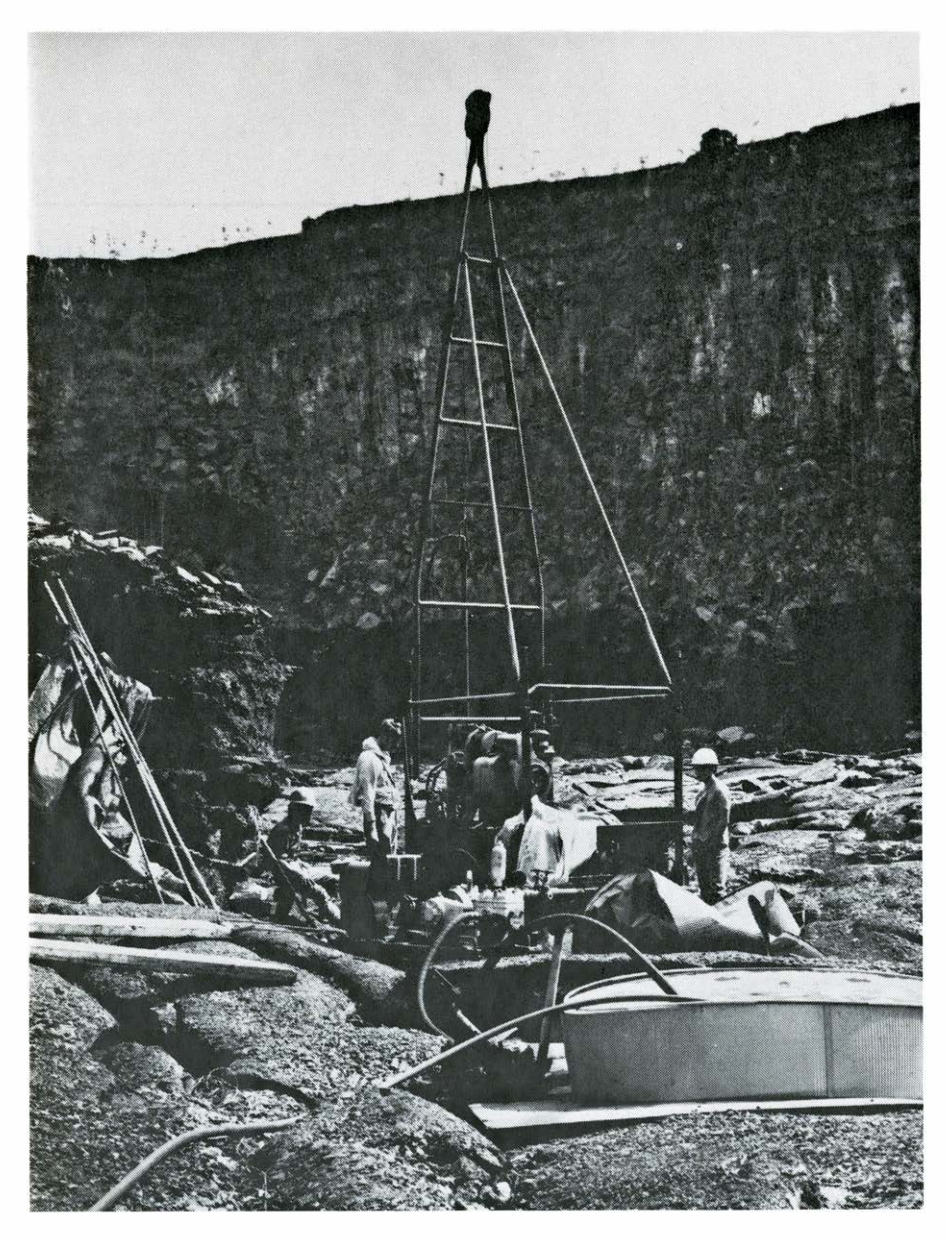
MAPA DE LA ZONA DE FRACTURA ORIENTAL del Kilauea; muestra la ubicación de los tres lagos de lava investigados por los autores: uno formado en 1959 en el cráter Kilauea Iki, otro en 1963 en el cráter Alae y el tercero en 1965 en el cráter Makaopuhi. El Halemaumau es la "hoya de fuego" de la caldera situada en la cima del volcán, cuya forma parece la de un plato invertido. En los últimos diez años, los flujos de lava nueva han cubierto los lagos Alae y Makaopuhi y parte de la carretera de los Cráteres (*señalada por líneas a trazos*) pero el lago Kilauea Iki no quedó cubierto.



BOLSON DE MAGMA situado a unos cuatro kilómetros por debajo de la cima del Kilauea; constituye la fuente de la lava que hizo erupción en superficie y se deslizó hacia las depresiones formando lagos de lava. El lago Kilauea Iki, de 120 metros de profundidad, se formó durante un proceso espectacular, en el que tuvieron lugar 17 erupciones en el transcurso de 36 días. El lago de lava Makaopuhi, cuya profundidad máxima es de 83 metros, se formó en una erupción que duró 18 días. El lago Alae, de 15 metros de profundidad, se originó en una erupción de tres días de duración.

expulsa, de la solución, el gas de la masa fundida. El gas escapa hacia la atmósfera o queda atrapado en la lava en forma de vesículas o burbujas. Las vesículas formadas durante la erupción de los tres lagos de lava del Kilauea, y que se enfriaron en la corteza a profundidades someras, suelen adquirir una configuración esférica cuyo diámetro puede llegar a ser hasta de un centímetro. Aparentemente se formaron a altas temperaturas, cuando la lava quedó sobresaturada de gases a causa de la reducción de la presión de confinamiento (por el mismo principio por el que aparecen burbujas en una botella de agua de soda cuando se quita el tapón). A mayor profundidad en el interior de la lava, la presión de confinamiento es también mayor, y, debido a ello, las vesículas resultan cada vez menores y más escasas. En el lago de lava Alae, por debajo de los seis metros de profundidad, la mayoría de las vesículas son diminutos poros angulosos de menos de un milímetro de diámetro. Al parecer, se formaron cuando se expulsó el gas de la solución a causa de la cristalización de la lava en enfriamiento. La composición de los gases expulsados de la lava también fue cambiando a medida que se enfriaba el lago: aumentaba el vapor de agua a expensas de los gases expulsados más rápidamente, correspondientes a compuestos de carbono y de azufre.

Cabría esperar que una lava solidificada se hundiera, ya que sus cristales son más densos que la masa fundida; sin embargo, la lava del lago Alae adquiría una mayor flotabilidad según se iba solidificando, pues estaba plagada de vesículas de gas. Debido a ello, a medida que la lente subsuperficial de lava fundida del centro del lago se solidificaba, empujaba hacia arriba la superficie de la corteza que la cubría. En los bordes del lago donde no había lente en proceso de solidificación porque las cortezas superior e inferior estaban fusionadas, la superficie subsidía a medida que la lava en enfriamiento se contraía térmicamente y se hacía cada vez más densa; como esta lava ya estaba solidificada, en ella no se formaban las burbujas que le hubieran dado mayor flotabilidad. También toda la superficie de los lagos Kilauea Iki y Makaopuhi experimentó un movimiento generalizado cuando los lagos se enfriaron. Como estos lagos eran más profundos que el lago Alae, las mayores presiones existentes en ellos impidieron la formación de la lava vesicular que hubiera empujado, hacia arriba, la superficie.



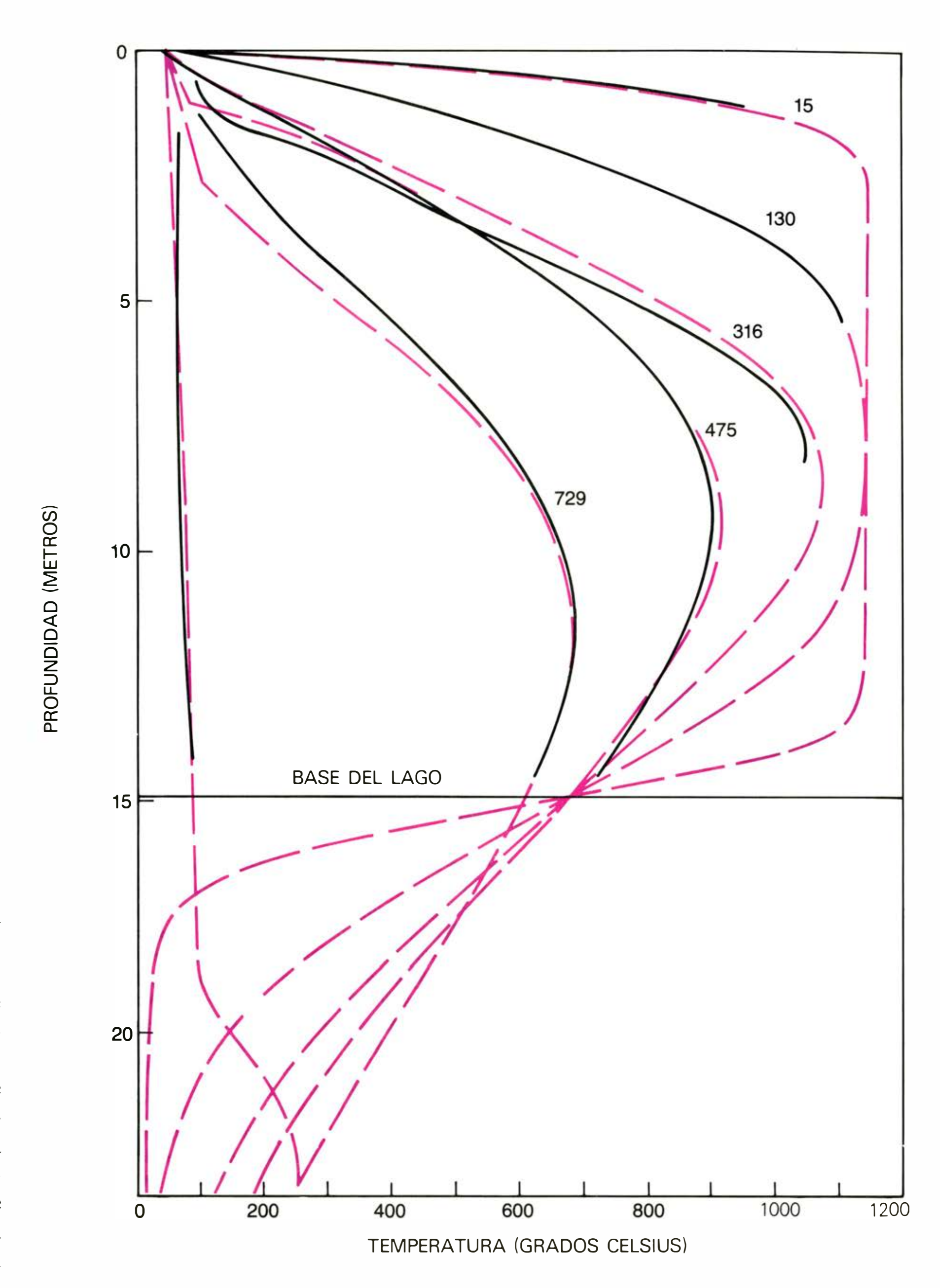
EQUIPO DE PERFORACION accionado por un motor de sierra en la superficie del lago de lava Makaopuhi. El aparejo, bajado a la superficie por un helicóptero, colectó muestras de lava sólida y fundida de distintas profundidades del lago. Frente al equipo de perforación hay un tanque de agua, para enfriar las muestras de lava candente. En la pared del cráter situada detrás del aparejo aflora un lago más antiguo de lava diaclasada en columnas, cubierto luego por varias coladas horizontales.

Investigamos la cristalización del basalto mediante la forma de muestras testigo de la lava semifundida situada en la base de la corteza superior, de la lava fundida subyacente (consecutiva a la anterior) y de la piedra pómez y lava fundida arrojadas por erupción a la superficie cuando se formaron los lagos. Las muestras nos permitieron reconstruir la historia de la cristalización de un lago de lava. Al principio del proceso los cristales embebidos en la lava fluida escasean, y entre ellos predominan los granos gruesos de olivino. Conforme disminuye la temperatura del lago, se van formando y se hacen cada vez más abundantes los cristales finos de piroxeno y los cristales claros, bandeados, de plagioclasa. En las proximidades de la base de la corteza, a 1070 grados C., los cristales de piroxeno y plagioclasa son mayores y más numerosos, constituyendo cerca del 50 por ciento de la lava. A 1075 grados los óxidos de hierro y los de hierro y titanio se encuentran todavía en estado vítreo, pero a 1065 grados ya pueden verse algunas láminas opacas de ilmenita (óxido de hierro y titanio). La abundancia de estos óxidos imprime al vidrio un vivo color pardo. La mayor parte de la lava enfriada a 1020 grados es cristalina: un mosaico de cristales mutuamente indentados de piroxeno, plagioclasa e ilmenita, rodeado de pequeñas cantidades de un vidrio pardo claro que contiene agujas de apatito y algunos diminutos cubos de magnetita. A 760 grados la lava está completamente solidificada pero no es del todo cristalina; mantiene algunas bolsas de un vidrio rico en sílice.

Los tipos de lava de los tres lagos tie-

nen una composición similar. Entre ellos existen diferencias químicas poco notorias, pero en cada una de ellas se aprecian variaciones químicas significativas. Tres procesos son los responsables de estas diferencias y variaciones. El primero es el hundimiento de los cristales de olivino magnésico, más densos que la masa basáltica fundida. Los cristales migran hacia abajo como granos de arena espolvoreados en un estanque. El efecto de la migración es especialmente patente en un antiguo lago solidificado que aflora en las paredes del cráter Makaopuhi, donde el olivino escasea en la superficie del lago, pero abunda a profundidades de seis a veinticuatro metros por encima del fondo. La mayor parte de las variaciones en la composición química de lavas constituidas por más del seis por ciento de óxido de magnesio son consecuencia de variaciones en la abundancia de olivino.

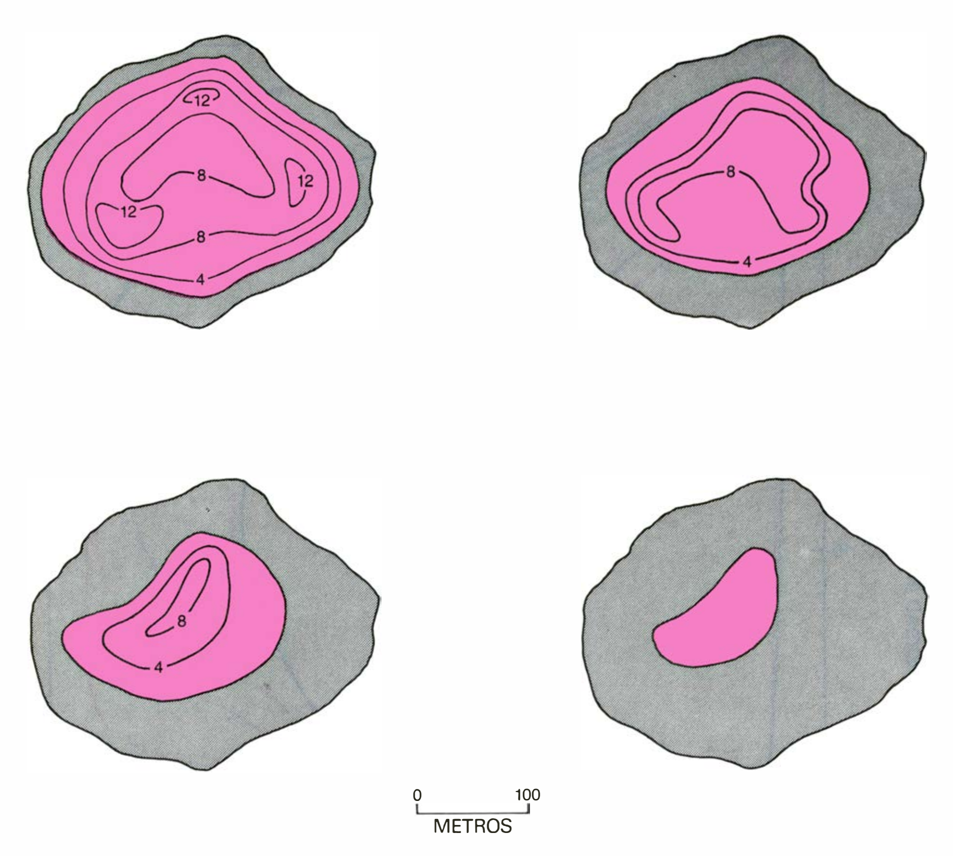
El segundo proceso, la separación de los cristales de plagioclasa y piroxeno en la lava fluida, resultante de diferencias de forma y densidad, da lugar a modificaciones más sutiles en la composición química. Advertimos por vez primera esta separación en el lago de lava Makaopuhi, cuando se observó que la lava fundida que había fluido hacia el interior del taladro contenía menor cantidad de cristales, en especial de cristales de piroxeno, que la lava fluida no canalizada. El tercer proceso, que entra en juego a temperaturas menores, consiste en la separación de la masa fundida intersticial del mosaico de cristales de la lava ya rígida, aunque no completamente cristalizada, próxima a la base de la corteza superior. Esta masa fundida rellena fracturas de la corteza y rezuma hacia los pozos de sondeo.



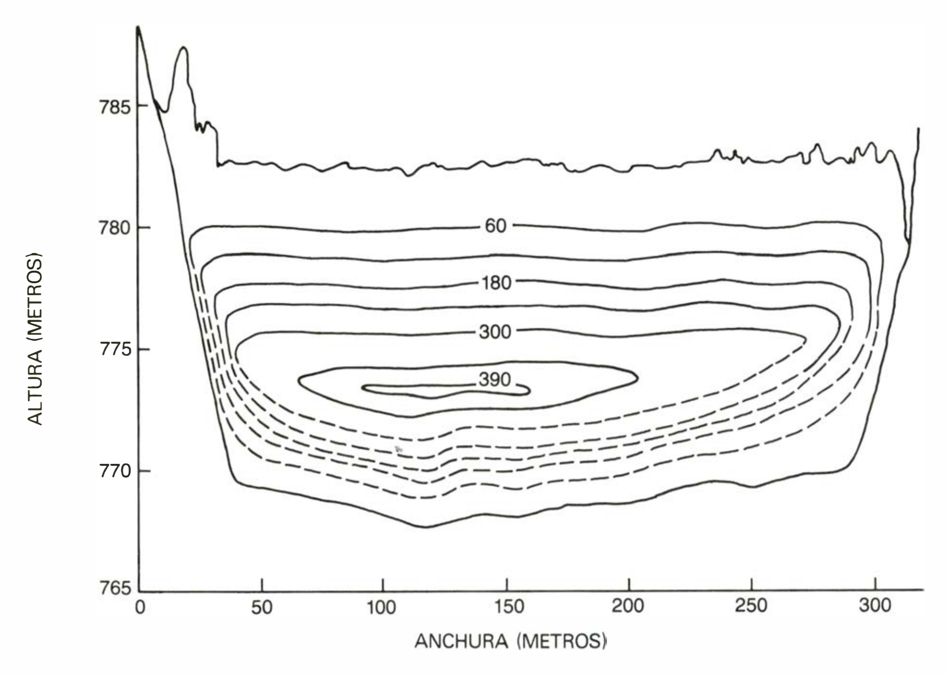
TEMPERATURA DEL LAGO DE LAVA ALAE en función de la profundidad. Las curvas continuas en negro se han construido en base a mediciones efectuadas con termopares revestidos de acero que se fueron introduciendo en los pozos de sondeo durante un período de cuatro años. En cada curva se indica el número de días transcurridos desde la formación del lago. Las curvas de color a trazos corresponden a un modelo numérico de la actividad térmica en un lago de lava deducido por ordenador. El modelo incluye los procesos externos que afectan al flujo de calor en un lago de lava (enfriamiento por la acción del aire y la roca circundantes, que inicialmente están a una temperatura de unos cero grados, y enfriamiento por el agua de lluvia), el calor latente del basalto (calor liberado por cristalización) y las variaciones de densidad, porosidad y conductividad térmica de la lava. La diferencia media entre las temperaturas calculada y medida no llega a los dos grados C. Esta notable concordancia entre lo previsto y lo obtenido tiende a confirmar la validez del modelo; los rasgos más importantes de este modelo son que la conductividad térmica varía con la porosidad y la temperatura de la lava y que la transformación en vapor del agua de lluvia que ha caído sobre la corteza caliente solidificada de un lago de lava desarrolla una preponderante influencia en la disminución de la temperatura del lago de lava.

La composición del basalto solidificado, pero todavía caliente, de la corteza superior de los lagos también resulta alterada por la acción de gases, formados en la lava subyacente en cristalización, que ascienden a través de las fracturas de la corteza. La alteración más sorprendente es la formación de películas rojas del óxido de hierro hematita en el olivino y la sustitución de minerales silicatados por óxidos a temperaturas comprendidas entre 500 y 800 grados C. Esta alteración se debe a variaciones en la concentración de oxígeno libre a diferentes profundidades de los lagos. Allí donde el oxígeno abunda, la forma más oxidada del hierro, es decir, la hematita, es sumamente estable. Donde el oxígeno abunda menos, es estable la forma menos oxidada del hierro, la magnetita. (La ausencia de oxígeno libre en la Luna explica que allí no haya ni hematita ni magnetita, el otro principal óxido de hierro. En la Luna es estable la forma menos oxidada del hierro, o sea, el hierro metálico.) A temperaturas comprendidas entre 550 y 750 grados, la concentración de oxígeno era elevada y la acumulación de hematita, muy notoria. Es de suponer que la elevada concentración de oxígeno fuera, en parte, consecuencia de la descomposición del vapor de agua contenido en los gases que ascendían por las fracturas.

Otro fenómeno que altera la corteza de enfriamiento de los lagos de lava es el desarrollo de grandes grietas en la misma a medida que se contrae el basalto que se está enfriando. Estas grietas se abren un minuto después, más o menos, de que la lava incandescente salga a la superficie en el transcurso de



BURBUJAS DE GAS que se forman en la lava que se está solidificando en el lago de lava Alae; imprimen un carácter abovedado a su superficie. Un lago de lava, al igual que otro de agua, es más somero en sus bordes que en el centro, de modo que la corteza superficial y la del fondo se unen en una sola en las orillas (*negro*), separando la lente fundida de la roca que limita la cuenca lacustre. La parte superficial del lago que cubre la lente está coloreada. Al enfriarse la lava, encerrada en el lago, el gas contenido en la masa fundida es expulsado de ella y empuja hacia arriba las zonas coloreadas de la superficie del lago. Las curvas de distribución indican la velocidad de elevación de la superficie en unidades de 10^{-2} centímetros por día. En los bordes del lago, donde ambas cortezas, superficial y profunda, se juntaban, y por tanto no había lente de material fundido en vías de solidificación, la superficie (*en negro*) se hundía a medida que la lava enfriada se iba haciendo cada vez más densa por contracción térmica; como esta lava ya se había solidificado, en ella no se formaban las burbujas que hubieran aumentado su flotabilidad. El diagrama de isohipsas de la izquierda representa el lago de lava tal como evolucionó en el período comprendido entre los 34 y los 104 días después de su formación; el de la derecha, entre los 236 y los 293 días; de los diagramas de abajo, el de la izquierda, entre los 293 y los 345 días, y el de la derecha, entre los 345 y los 400 días. (Ilustración de Walken Graphics.)



CORTE LONGITUDINAL del lago de lava Alae; muestra cómo la isoterma (superficie de igual temperatura) de 1000 grados C. cambiaba a medida que se iba reduciendo la lente de lava fundida. En cada posición de la isoterma se indica cuántos días habían transcurrido desde que se formó el lago.

una erupción. Conforme sigue enfriándose y engrosándose la corteza, las grietas se propagan hacia abajo por nueva fracturación y se abren otras grietas que dividen la superficie de la corteza en polígonos. En pocas horas, los polígonos quedan deformados al sufrir sus centros un abombamiento hacia arriba debido a la expansión vesicular de la lava subyacente que se está solidificando. Se abre luego un número todavía mayor de grietas que, posteriormente, se caracterizarán porque en ellas los gases ascendentes habrán depositado azufre y yeso. Una cartografía reiterada de las grietas superficiales del lago de lava Alae indica que el ritmo de fracturación es mayor cuando la lluvia enfría la corteza.

Los lagos de lava del Kilauea ofrecen una buena oportunidad para el estudio de las propiedades físicas del magma basáltico. Herbert R. Shaw, del Servicio Geológico de los Estados Unidos, y dos de nosotros (Peck y Wright) medimos la viscosidad de la lava fluida, tanto en el laboratorio como en el lago de lava Makaopuhi. La viscosidad se define como el grado de resistencia a fluir que tiene un fluido cuando se halla sometido a una fuerza. Colocamos en el pozo de sondeo, en la masa fundida, a profundidades comprendidas entre 2,6 y 3,2 metros, un viscosímetro de acero inoxidable, consistente en un rotor que mide la viscosidad por la resistencia al giro. Allí la lava tenía una temperatura de 1130 a 1135 grados C. y contenía un 25 por ciento de cristales y de un dos a un cinco por ciento de burbujas gaseosas. Nuestros resultados confirmaron los obtenidos en el laboratorio, donde la viscosidad subió desde 100 poise a 1300 grados hasta unos 320 poise a 1200 grados. (Un poise equivale a una dina por segundo por centímetro cuadrado.) Este cambio es comparable a la diferencia de viscosidad entre un aceite pesado y una melaza espesa a temperatura ambiente. A temperaturas inferiores a 1200 grados la viscosidad aumenta bruscamente debido a la acumulación de cristales en la masa fundida.

La presencia de burbujas gaseosas a estas temperaturas hace que la lava sea más plástica que fluida. El comportamiento fluido se caracteriza por una respuesta tal que se produce deformación por pequeño que sea el esfuerzo aplicado, mientras que en el comportamiento plástico sólo se produce deformación si el esfuerzo aplicado rebasa un mínimo determinado. A velocidades bajas de rotación del viscosímetro, la lava a 1130 grados tenía una viscosidad de 7500 poise. Las burbujas de gas "espesaban" la lava del mismo modo

que el aire espesa las claras de huevo cuando se las bate. A velocidades de rotación mayores, las burbujas se alargaban, formando planos de debilidad que disminuían la viscosidad hasta los 6500 poise.

Las investigaciones de los lagos de lava en los años sesenta se emprendieron para conocer mejor la naturaleza del magma. En los últimos cuatro años, se ha intensificado el estudio de los lagos de lava como parte de la investigación para encontrar fuentes de energía alternativas. Se considera verosímil que el magma del subsuelo de los Estados Unidos pueda proporcionar anualmente 13×10^{21} calorías, cantidad equivalente a la que se liberaría en la combustión de nueve billones de barriles de petróleo.

Los obstáculos interpuestos para la explotación de esta energía son muchos. Se tendrán que desarrollar métodos para localizar masas magmáticas y determinar su tamaño, forma, profundidad y contenido calorífico. Se requerirán técnicas especiales para perforar un depósito magmático situado a gran profundidad, para instalar en él algún mecanismo capaz de transferir el calor y para extraer la energía por el pozo de sondeo. El aprovechamiento de la energía magmática también plantea un reto a los científicos que investigan materiales: habrá que buscar materiales capaces de operar satisfactoriamente durante períodos largos en el ambiente caliente, corrosivo y de gran presión que es un magma.

El lago de lava Kilauea Iki constituye un laboratorio natural para llevar a cabo investigaciones encaminadas a superar estos obstáculos. La lava del lago sufrió una degasificación parcial durante la erupción y se encuentra a una presión relativamente baja, debido a lo cual no puede tomarse el lago como un arquetipo válido de un magma de profundidad. Sin embargo, el magma arrojado a la superficie es mucho más accesible y, gracias a ello, hemos podido utilizar los conocimientos adquiridos sobre las propiedades físicas y químicas de la masa fundida del lago aplicándolos a ensayos de posibles técnicas para deducir la posición, el tamaño y la forma de masas magmáticas de profundidad. Nuestra investigación del lago es totalmente experimental, dirigida a rellenar determinados huecos del saber actual y no a aprovechar ningún lago concreto de lava del Kilauea como fuente energética.

En 1976 se hicieron dos sondeos de la corteza solidificada del lago hasta el límite superior de la lente de lava fundida. Las propiedades de la lente se determinaron directamente en base a muestras testigo y a instrumental introducido a través de los pozos. Tratamos también de determinar indirectamente las mismas propiedades mediante técnicas tales como mediciones sísmicas y electromagnéticas efectuadas en la superficie del lago. En principio estas técnicas podrían aplicarse a depósitos profundos de magma, inaccesibles por sondeo directo. Un análisis de los resultados sugiere que las técnicas sísmicas y electromagnéticas permiten determinar los límites superior y laterales de la lente de lava fundida, pero no dejan evaluar con precisión su grosor. Al parecer, estas técnicas no identifican como fundida toda la lente de lava líquida, sino sólo la parte central, más caliente.

Bernard Chouet, del Instituto de Tecnología de Massachusetts, y John L. Colp, de los Laboratorios Sandia, han cartografiado, con la ayuda de sismómetros, los límites aproximados de la lente fundida del lago de lava Kilauea Iki. La energía sísmica la proporcionan débiles temblores provocados por agrietamientos de contracción térmica en las regiones que tienen fuertes gradientes térmicos en los bordes superior y laterales de la lente. Llegan a presentarse hasta 200 sacudidas por hora. Los estudios electromagnéticos del lago hechos por Lennart A. Anderson y Charles J. Zablocki, del Servicio Geológico de los Estados Unidos, indican unos límites muy similares: una lente de lava fundida de 500 metros de grosor y 750 metros de longitud.

Los estudios electromagnéticos señalan variaciones de la conductividad eléctrica a diferentes profundidades del lago. La capa más externa de la corteza superior, cuya temperatura no alcanza los 100 grados C., es muy buena conductora, debido a que las grandes grietas que la dividen están rellenas de agua de lluvia. Bajo ella hay una capa de basalto solidificado que, a temperaturas superiores a los 100 grados C., ofrecen una gran resistividad, lo cual se debe tanto a la ausencia de una capa líquida como a que los iones teóricamente conductores que constituyen el basalto están inmovilizados en forma de minerales. A profundidades de más de 40 metros la lente de lava se encuentra fundida; es una excelente conductora porque sus iones tienen libertad de movimiento.

A lo largo del curso presente se han hecho seis nuevos sondeos en el lago de lava Kilauea Iki. La lente de lava parcialmente fundida, que mide unos 30 metros de espesor, tiene ya muchos cristales. Se solidificará por completo en un lapso de cinco años. Hasta que el Kilauea o cualquier otro volcán no dé origen a un nuevo lago de lava esta oportunidad única de estudiar el magma *in situ* será muy pronto cosa del pasado.

Evolución de los Crocodilios

Unos 200 millones de años después de que estos reptiles aparecieran y se diversificaran, perduran tres familias: los aligátores, los cocodrilos y los gaviales. Son los últimos arcosaurios, los reptiles dominantes del Mesozoico

Eric Buffetaut

A lo largo de los últimos 200 millones de años un orden de reptiles, los Crocodilios, ha demostrado repetidamente una de las principales características de la evolución: si existe un determinado tipo de nicho ecológico, es probable que lo ocupe un tipo de organismo determinado. Por ejemplo, existen tres familias actuales de crocodilios: los aligátores y caimanes (Aligatóridos), los cocodrilos (Crocodilios) y los gaviales (Gaviálidos). Los caimanes tienen la cabeza roma, los cocodrilos la tienen más alargada y los gaviales extremadamente estrecha. Las mismas adaptaciones evolutivas de cabezas alargadas y redondeadas surgieron repetidamente en los crocodilios que hace mucho tiempo que se extinguieron y que estaban emparentados sólo lejanamente con los que viven hoy en día, si es que lo estaban. Es evidente que el nicho ecológico que favorece la evolución de reptiles de hocico largo que se alimentan de peces ha sido ocupado de forma repetida por los adaptables crocodilios. Otro ejemplo: los crocodilios actuales se hallan por lo general adaptados a una vida semiacuática en un ambiente cálido. Sin embargo, en el pasado algunos miembros del grupo se adaptaron a nichos situados en tierra firme que en la actualidad están ocupados por mamíferos carnívoros. Al mismo tiempo, otros se hicieron completamente acuáticos, con una cola parecida a la de los tiburones y los miembros remodelados en aletas nadadoras.

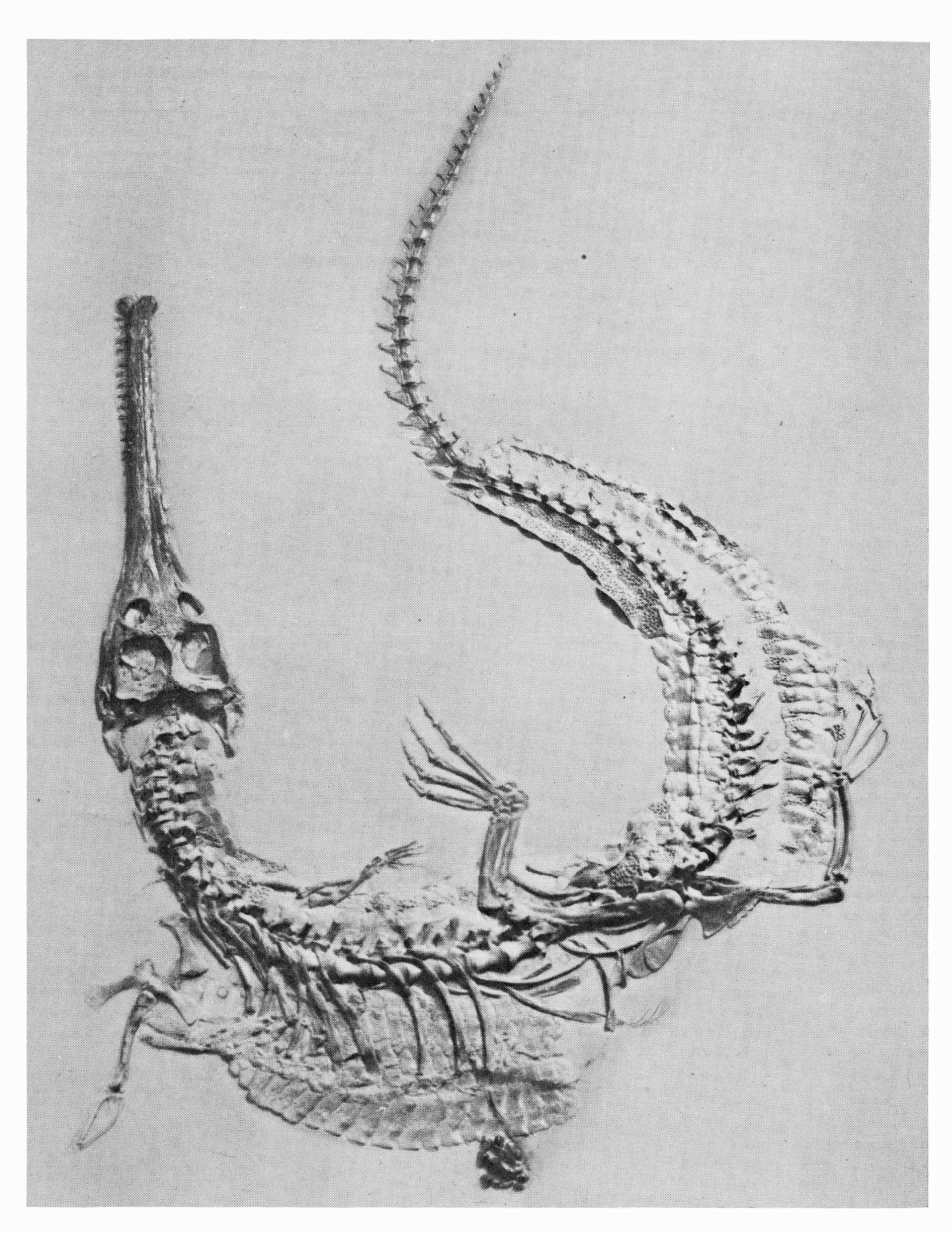
Voy a explicar aquí lo que se sabe acerca de la evolución de este notable grupo de reptiles desde su aparición en el registro fósil hacia finales del Triásico. En este período, los reptiles eran las formas dominantes de la vida animal en nuestro planeta, ocupando la tierra, el mar y el aire. Cuando la paleontología científica llegó a su mayoría de edad a principios del siglo XIX, entre los primeros animales vertebrados que se exhumaron estaban los cocodrilos fósiles, y aún hoy en día se descubren con cierta frecuencia nuevos crocodilios fósiles. En consecuencia, y aunque sigue habiendo lagunas en su historia, resulta evidente que el orden crocodilios se diversificó considerablemente a lo largo de los últimos cientos de millones de años, extendiéndose a todas las regiones del mundo y adaptándose a una gran variedad de nichos ecológicos.

Ya en 1875 T. H. Huxley, amigo y defensor de Darwin, reconocía que algunos crocodilios fósiles eran claramente distintos de otros, y también de los crocodilios vivos, con respecto a dos características anatómicas. La primera de estas características es la situación de los orificios nasales internos: las aberturas internas del conducto nasal que deja pasar el aire a través de la faringe hasta los pulmones. En los crocodilios actuales un largo septo óseo, el paladar secundario, divide la cavidad oral desde el conducto nasal. De este modo, los orificios nasales externos, situados cerca del extremo del hocico, están separados de los internos, cerca de la parte posterior del cráneo, detrás del paladar secundario, por una distancia sustancial. En los primeros crocodilios fósiles esta distancia era mucho menor.

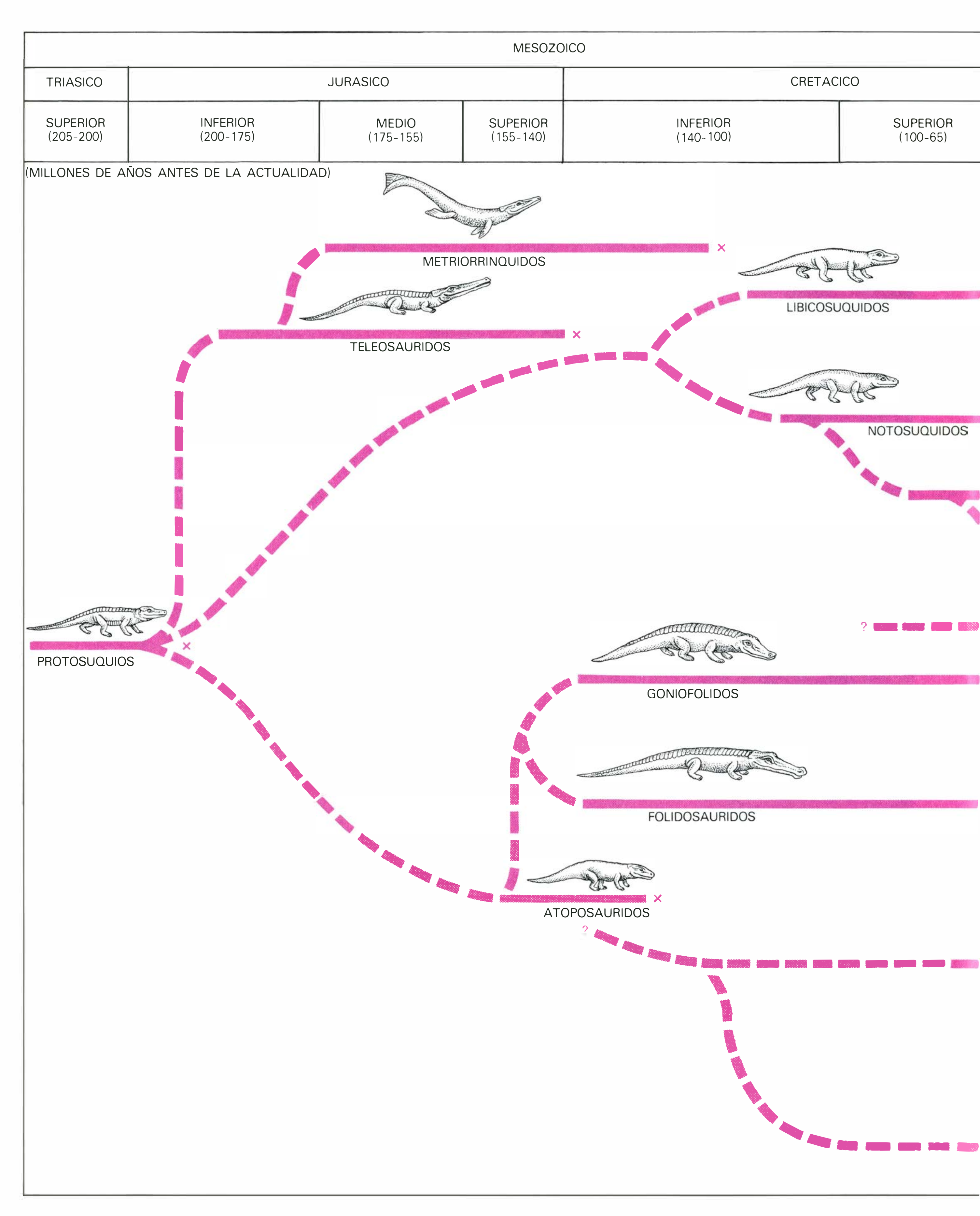
La segunda característica anatómica es la forma de las vértebras. En los crocodilios actuales la articulación de la columna vertebral es una disposición en rótula: la parte posterior redondeada de cada vértebra encaja en una cavidad de la parte anterior de la vértebra situada detrás. Como ocurre con la ubicación retrasada de los orificios nasales internos, no era éste el caso en los crocodilios primitivos. Huxley propuso que hubo un progreso desde una anatomía primitiva hasta otra más evolucionada con respecto a las dos características, y los descubrimientos fósiles desde su época han confirmado la validez general de su observación.

Los Crocodilios pertenecen a la misma subclase de reptiles que incluye los Dinosaurios y los Pterosaurios voladores: los Arcosaurios ("reptiles dominantes"). Los crocodilios más antiguos que se conocen, miembros del suborden Protosuquios, se encuentran como fósiles en formaciones rocosas continentales de finales del Triásico en Norte y Sudamérica (Arizona y Argentina) y en Africa (Sudáfrica y Lesotho). Ejemplares de reptiles fósiles de Inglaterra y de China podrían pertenecer asimismo a este suborden del Triásico tardío. Los Protosuquios eran animales pequeños, que raramente sobrepasaban un metro de longitud; su hocico era bastante corto y sus patas bastante largas. Una fuerte armadura de placas óseas cubría su dorso y vientre. Su presencia en rocas continentales sugiere que algunos de ellos eran carnívoros terrestres, pero entre los distintos géneros de Protosuquios hay otros que ya muestran adaptaciones a una vida acuática. La distribución geográfica aparentemente amplia de los Protosuquios, sobre todo si se incluyen Inglaterra y China, es en parte ilusoria. En el Triásico los continentes actuales todavía se hallaban unidos en una masa continental única y las barreras a la migración eran pocas.

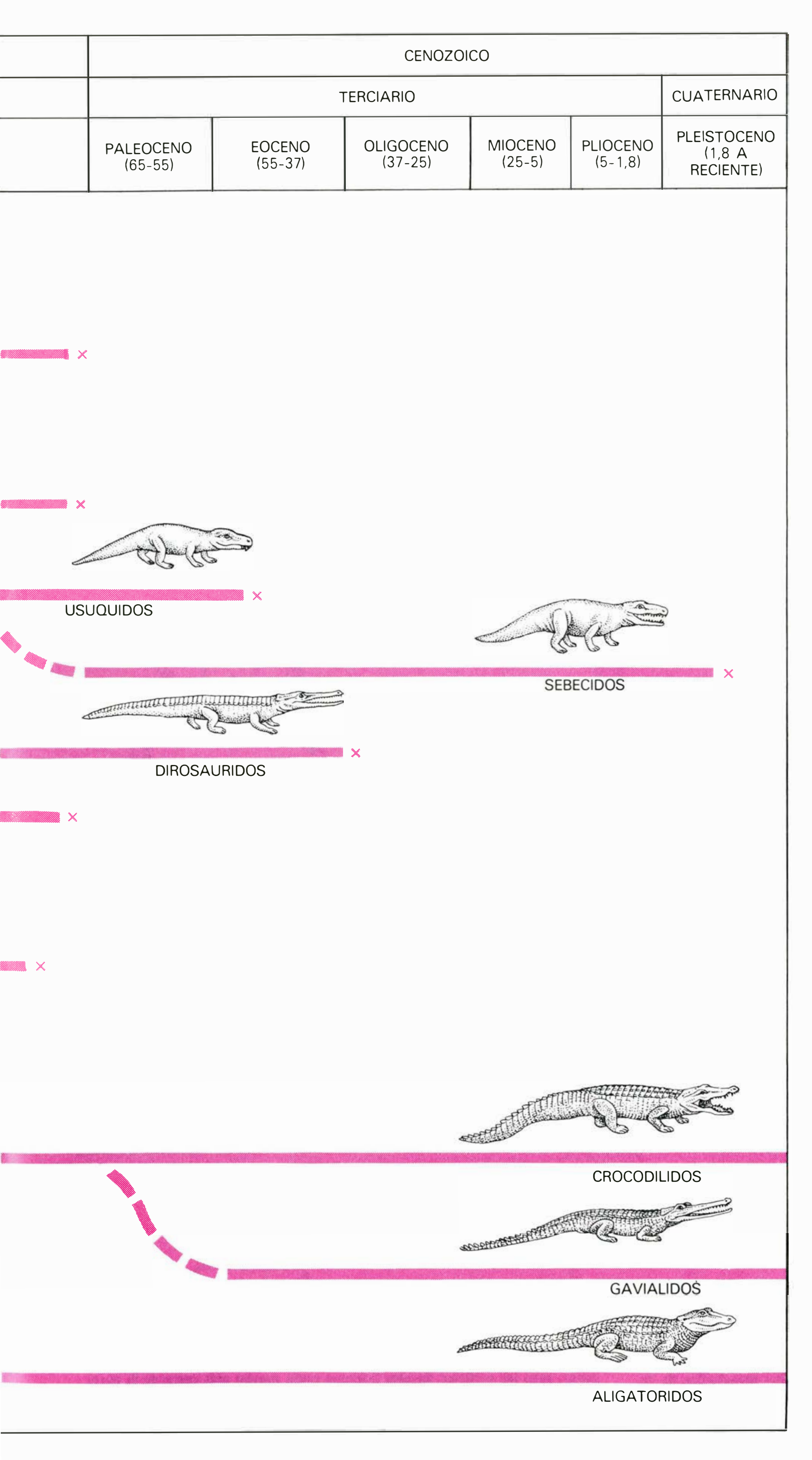
Todos los Protosuquios exhiben rasgos anatómicos que recuerdan los de los reptiles primitivos. Uno de ellos es la forma de las vértebras que Huxley había señalado: cada vértebra es ligeramente cóncava en ambos extremos. Otro es la presencia de aberturas pares a los lados del hocico, delante de las órbitas oculares. La función de estas aberturas no está clara. Finalmente, el paladar secundario está formado por menos huesos que en los crocodilios posteriores. No obstante, el número de géneros de Protosuquios y su diferenciación sugieren que hacia finales del Triásico ya hacía algún tiempo que habían ido evolucionando a partir de una estirpe ancestral.



CROCODILIO FOSIL del Jurásico inferior, encontrado en una formación de esquistos en Ohmden, en el sudoeste de Alemania. El ejemplar se encuentra en la colección del Instituto Paleontológico de la Universidad de Tubinga. Este animal de tres metros pertenece al género *Steneosaurus*, de la familia Teleosáuridos, adaptada a la vida en los mares someros que invadieron los bordes de los continentes en el Jurásico, período que empezó hace unos 200 millones de años. Superficialmente, los Teleosáuridos se parecen a los gaviales de Asia, de hocico estrecho, una de las tres familias de cocodilios actuales. Sin embargo, no están emparentados. Los Teleosáuridos se extinguieron a principios del Cretácico, hace unos 130 millones de años.



EVOLUCION DE LOS CROCODILIOS durante un período de más de 200 millones de años, trazada a lo largo de sus ramas principales (*líneas de color*). Estas ramas están conectadas de manera supuesta (*líneas de trazos*) o insegura (*interrogantes*). En el punto de partida se encuentra el suborden Protosuquios, que se halla en formaciones fósiles del Triásico superior. Algunos de estos pequeños animales terrestres exhiben una adaptación a la vida acuática. Una laguna separa a estos primeros crocodilios de los del suborden Mesosuquios, constituido por grupos que aparecen en formacio-



nes fósiles del Jurásico y del Cretácico. Sólo unas cuantas familias de mesosuquios persistieron hasta entrado el Cenozoico; los últimos miembros del suborden pudieron haber sobrevivido en Australia hasta hace alrededor de un millón de años. El suborden Eusuquios apareció durante el Cretácico, pero los gaviales no surgieron de los cocodrilos propiamente dichos hasta principios del Terciario (*abajo, derecha*).

Después del final del Triásico existe una laguna en el registro fósil de los Crocodilios que ocupa varios millones de años. Hasta finales del Jurásico inferior no aparecen ejemplares completos, y son muy diferentes de los Protosuquios del Triásico. Prácticamente todos los crocodilios de los yacimientos fósiles del Jurásico inferior y medio se encuentran en rocas marinas, no en las continentales, rocas que fueron depositadas en los mares someros que por aquella época habían invadido grandes zonas continentales.

Los crocodilios jurásicos más antiguos, miembros del suborden Mesosuquios, se hallaban entre los primeros fósiles vertebrados que la ciencia conoció. Eran representantes de la familia Teleosáuridos, y se encontraron con cierta abundancia en formaciones rocosas de Inglaterra, Francia y Alemania, ya en el siglo XVIII. Los Teleosáuridos, algunos de ellos pequeños pero otros de hasta cuatro metros o más de longitud total, recuerdan superficialmente a los gaviales actuales; es decir, tenían un hocico largo y mandíbulas estrechas con numerosos dientes finos. Debieron ser eficaces depredadores de peces.

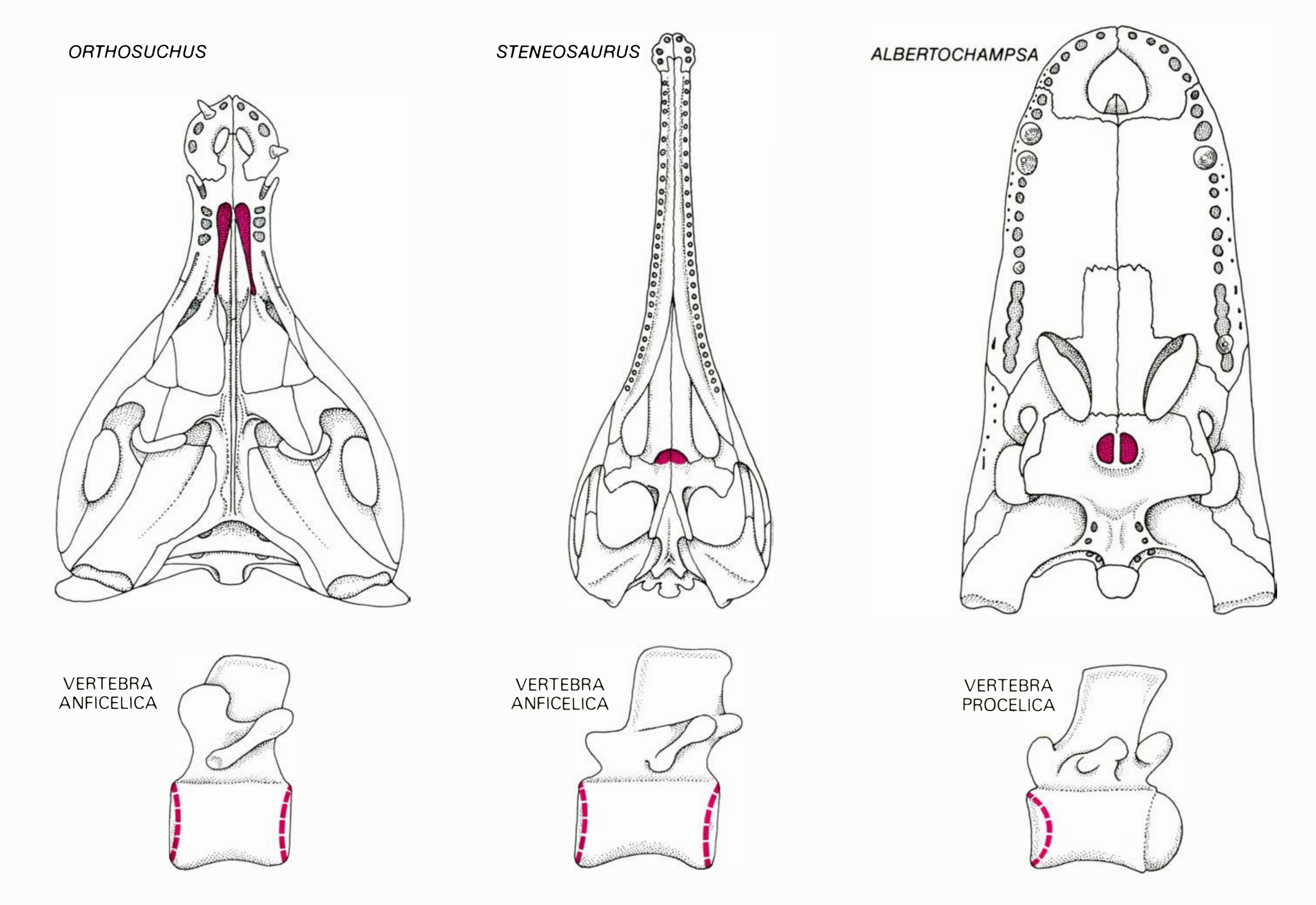
Con respecto a las dos características anatómicas primitivas que caracterizaban a los Protosuquios del Jurásico, los Mesosuquios ocupan una posición intermedia: sus vértebras eran todavía cóncavas en ambos extremos, pero sus orificios nasales internos estaban más cerca de la parte posterior del cráneo, y en el paladar había huesos adicionales. Su cuerpo estaba todavía protegido por una coraza de placas óseas, pero sus miembros, en particular los anteriores, eran cortos. Efectivamente, estos crocodilios estaban adaptados a ocupar el nicho de un vertebrado marino piscívoro. Aparentemente, los teleosáuridos se hallaban distribuidos por todo el mundo; sus restos se encuentran no sólo en Europa, sino también en Asia, Africa del Norte, Madagascar y Norte y Sudamérica. Prosperaron en Europa hasta el final del Jurásico, pero se extinguieron poco después.

Durante el Jurásico inferior apareció un segundo grupo de crocodilios marinos, probablemente a partir de antepasados teleosáuridos. Se trata de la familia Metriorrínquidos. Son los miembros más especializados en todo el orden Crocodilios y se hallan profundamente modificados para la vida acuática. Su parentesco con los Teleosáuridos es evidente por sus vértebras doblemente cóncavas, la estructura de su paladar y otras características de su anatomía craneal.

Fuera de eso, no obstante, el cráneo es fusiforme, las extremidades están transformadas en aletas y el extremo de la columna vertebral se halla fuertemente inclinada hacia abajo. En algunos yacimientos fósiles del Jurásico superior de Alemania se han conservado las impresiones de la piel de metriorrínquidos; estas impresiones muestran que la cola doblada de tales animales portaba una aleta caudal como la de los tiburones. La armadura de placas óseas característica de los Teleosáuridos y de sus predecesores triásicos ha desaparecido por completo.

Los Metriorrínquidos debieron ser buenos nadadores; los estudios de las rocas que contienen sus restos indican que, a diferencia de sus primos de aguas someras, se aventuraban a grandes distancias de la costa. Se especula sobre su forma de reproducirse. Pudieran haber parido crías vivas, como hoy en día hacen ciertas serpientes de mar, o bien quizás emularan a las tortugas marinas de nuestros días, arrastrándose a tierra para poner sus huevos. Sus restos fósiles se han encontrado sobre todo en Europa, pero Z. B. de Gasparini, del Museo de Historia Natural de La Plata, en Argentina, y sus colaboradores han descrito recientemente cráneos bien conservados de metriorrínquidos exhumados en formaciones jurásicas de Argentina y Chile. Los metriorrínquidos de aguas profundas se extinguieron a principios del Cretácico; no fueron sustituidos por ningún crocodilio de adaptaciones similares.

El final del Jurásico estuvo marcado por una retirada general de los mares someros que habían invadido los continentes del mundo. Una consecuencia de este retroceso es que las formaciones rocosas que contienen fósiles de animales terrestres y dulceacuícolas del Jurásico superior y del Cretácico inferior son relativamente abundantes. Entre estos vertebrados de la última parte del Mesozoico hay varios crocodilios que estaban adaptados a ambientes terrestres, dulceacuícolas y costeros. Su aparición súbita en el registro fósil en esta época indica un largo período de evolución y diversificación previas, pero en la mayor parte de los casos sus relaciones con los crocodilios anteriores permanecen oscuras. Quizá los nuevos crocodilios habían surgido muy tierra adentro de los continentes; la ausencia de yacimientos rocosos continentales que se formaran durante el Jurásico inferior y medio hace difícil reconstruir el pasado fósil de estos animales. Unos de estos nuevos crocodilios, la familia Atoposáuridos, se hallan por suerte bien conservados en las calizas litográficas de Europa. Un estudio detallado de sus esqueletos, notablemente completos, realizado por P. Wellnhofer, de la Colección del Estado Bávaro para la Paleontología y la



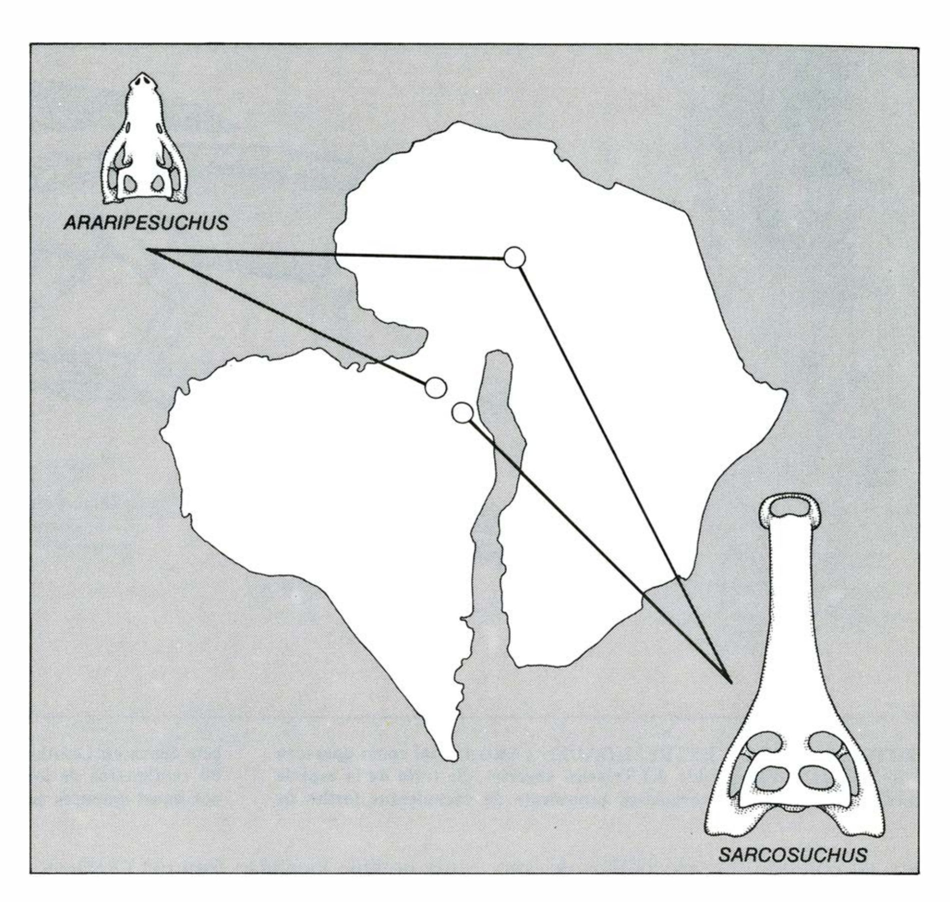
CRANEOS Y VERTEBRAS indican los tres estadios principales de la evolución de los Crocodilios. En los Protosuquios típicos, ejemplificados a la izquierda por una especie del género *Orthosuchus*, de una formación del Triásico superior en Africa, los orificios nasales internos (*color*) se encuentran cerca de la parte anterior del cráneo y las vértebras son doblemente cóncavas (*líneas de trazos de color*). En el mesosuquio del centro, una especie del género teleosáurido *Steneosaurus*, parecido a un gavial, de una formación del Jurásico en Europa, las vértebras son todavía doblemente cóncavas, pero los orificios nasales internos están mucho más cerca de la parte posterior del cráneo. En el eusuquio de la derecha, una especie del género de aligatóridos *Albertochampsa*, de una formación del Cretácico superior en Canadá, se ha desarrollado ya la condición moderna: los orificios nasales internos se sitúan aún más hacia atrás, abriéndose en los huesos pterigoides del paladar, y las vértebras son cóncavas-convexas. Los ejemplares no se han dibujado a la misma escala: *Orthosuchus* era pequeño y los demás, mayores. (Los eusuquios abarcan las tres familias de crocodilios actuales.)

Geología Histórica, de Munich, indica que estos pequeños crocodilios, de no más de 40 centímetros de longitud por término medio, eran probablemente ágiles carnívoros terrestres.

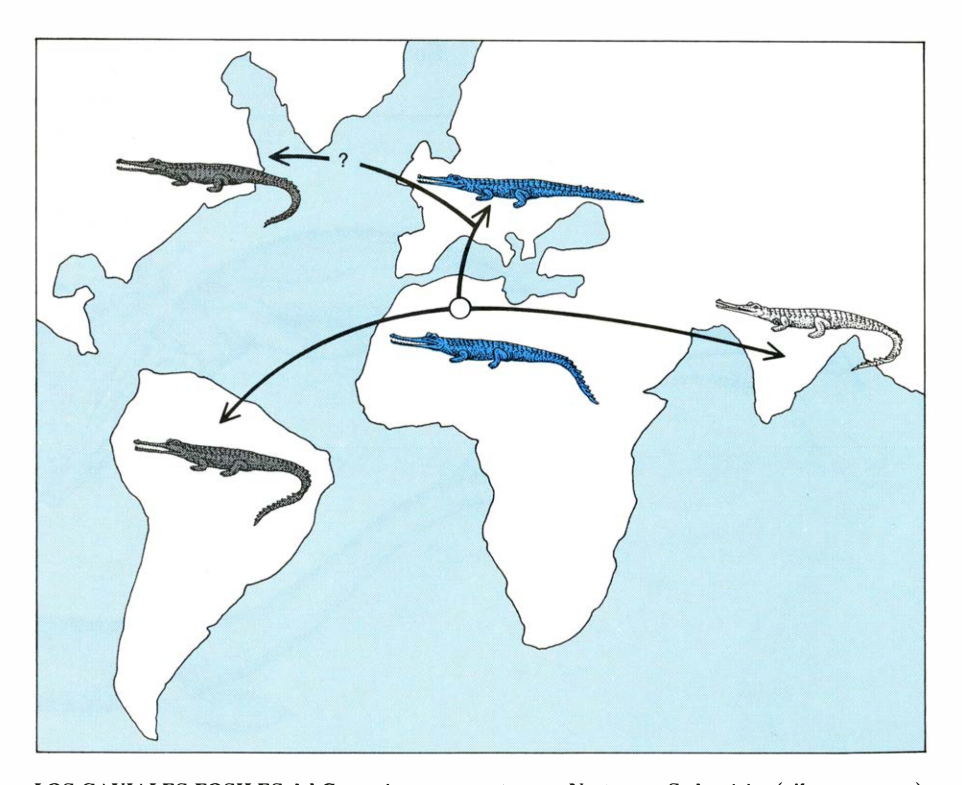
Una segunda familia de recién llegados, los Goniofólidos, eran muy parecidos a los caimanes y cocodrilos de hoy día por su aspecto general y su modo de vida. Eran animales bastante grandes, con un hocico ancho y moderadamente largo, y fuertes dientes. Sus vértebras eran del tipo primitivo, doblemente cóncavas, y el techo de su paladar, aunque era algo más evolucionado que el de los crocodilios primitivos, seguía siendo de estructura mesosuquia. El nicho ecológico que ocuparon los Goniofólidos debió ser similar al que explotan los cocodrilos actuales. Sus fósiles se encuentran en formaciones del Jurásico superior y del Cretácico inferior, tanto en Europa occidental como en Norteamérica. La distribución no ha de sorprendernos. Por aquella época, el Atlántico todavía no se había abierto por completo y los animales podían desplazarse libremente entre Norteamérica y Eurasia.

En estas mismas formaciones fósiles se han mantenido conservados otros crocodilios, la mayoría de ellos pequeños. Algunos habían desarrollado ya vértebras de articulación en rótula y orificios nasales internos retrasados, características típicamente modernas, pero estos rasgos suelen hallarse combinados con características primitivas, de manera que es difícil clasificar a estos animales de un modo satisfactorio. Sólo puede decirse, generalizando, que a comienzos del Cretácico varias estirpes de pequeños mesosuquios mostraban una tendencia hacia la adquisición de rasgos progresivos. Los Eusuquios, con toda probabilidad, descendían de una de estas estirpes; se trata del suborden que incluye a los crocodilios vivos. Sin embargo, es difícil decidir a qué estirpe en concreto pertenecen. Los candidatos posibles son *Theriosuchus*, un género de los yacimientos fosilíferos de Weald, en Inglaterra, y una forma emparentada de Norteamérica que está estudiando Wann Langston, Jr., de la Universidad de Texas en Austin.

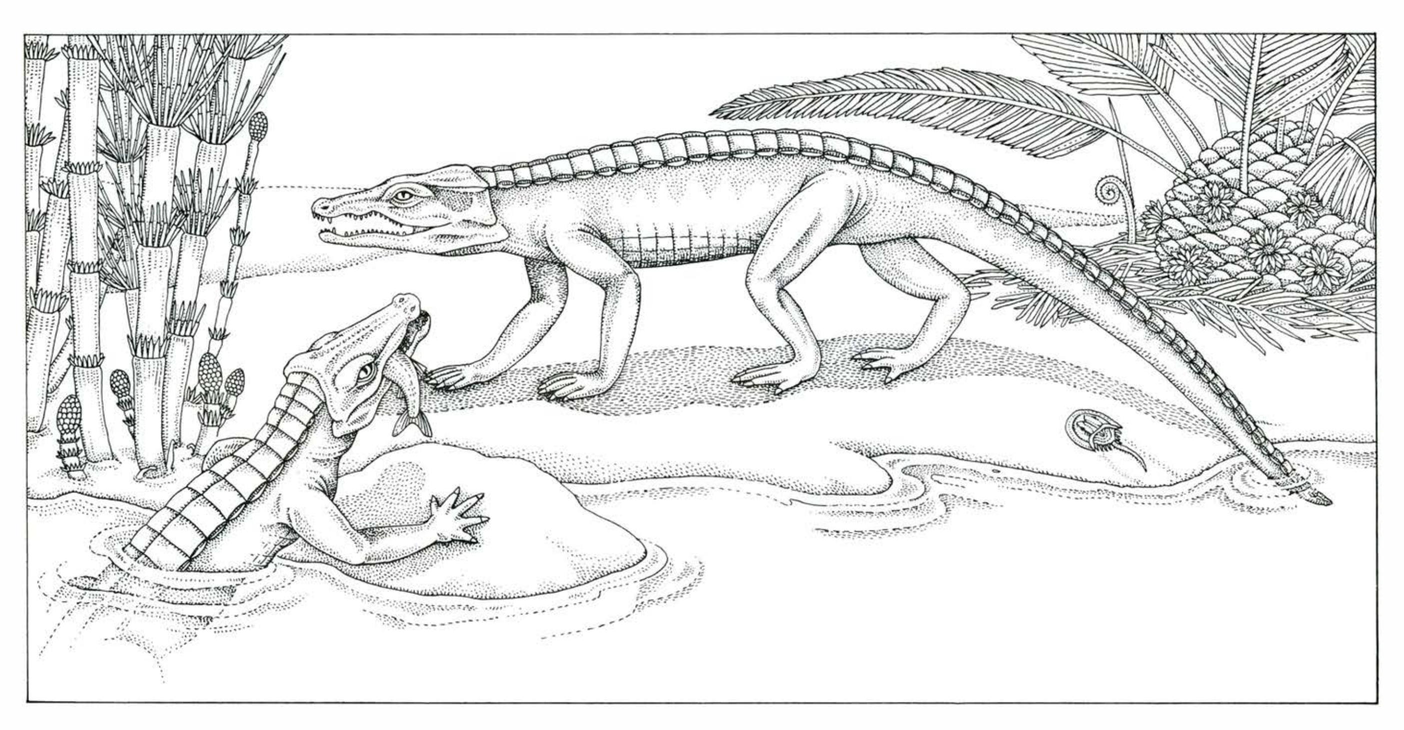
En el bloque austral de continentes, conocido como Gondwana, la situación a comienzos del Cretácico era algo diversa. Los diferentes componentes de Gondwana compartían unos cuantos grupos de crocodilios con el bloque continental septentrional, llamado Laurasia. Entre ellos se encontraban los Folidosáuridos, de hocico largo,



LOS MISMOS CROCODILIOS del período Cretácico inferior se encuentran en Sudamérica y en Africa, que entonces todavía estaban conectadas. Uno es el folidosáurido dulceacuícola gigante *Sarcosuchus*, que se reconoció por primera vez en una formación fósil del Níger. El otro es *Araripesuchus*, un mesosuquio terrestre de hocico corto descubierto en el Brasil y también en los sedimentos nigerianos.



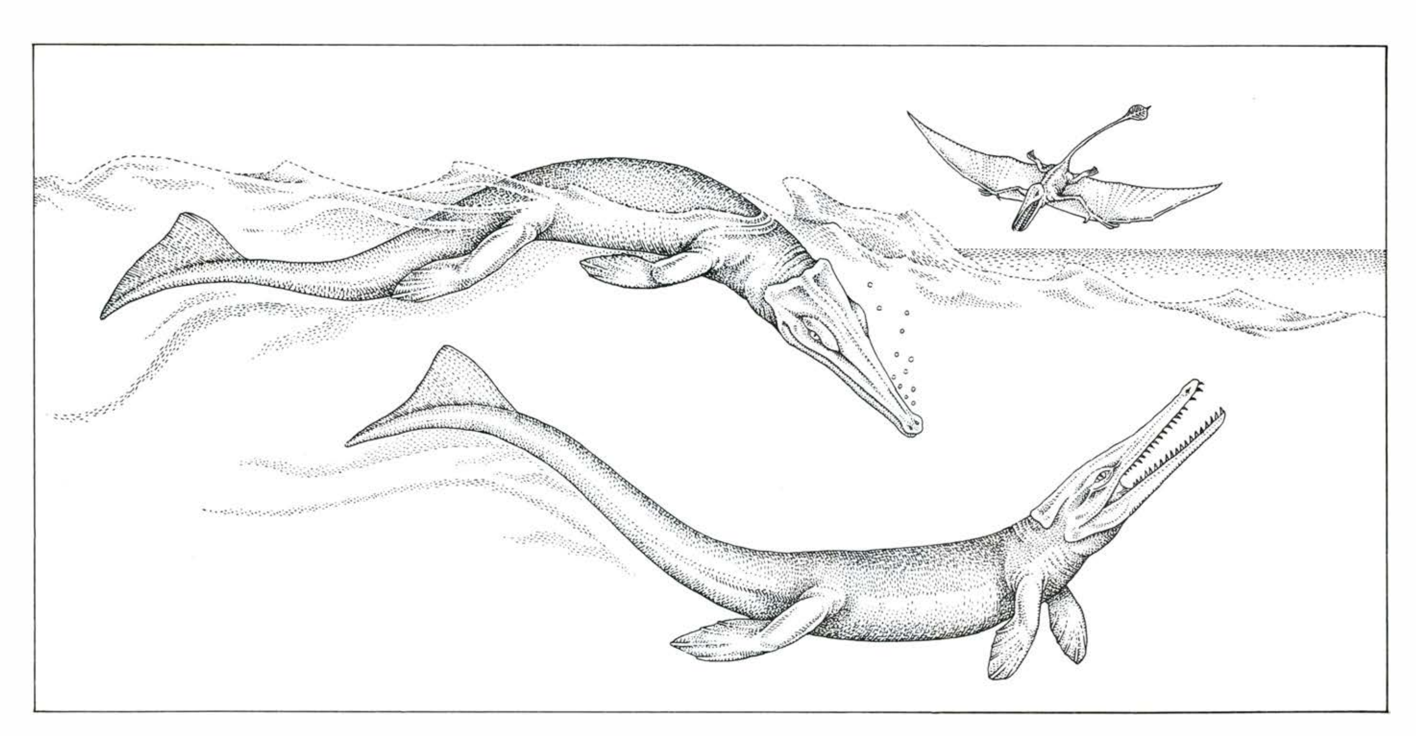
LOS GAVIALES FOSILES del Cenozoico se encuentran en Norte y en Sudamérica (*siluetas oscuras*), pero los gaviales actuales (*silueta clara*) están confinados a los grandes ríos del Paquistán, la India y Birmania. Ello era difícil de explicar hasta que ciertos crocodilios de hocico largo del Cenozoico temprano en Africa del Norte y Europa (*siluetas de color*) se identificaron como gaviales primitivos. Los migrantes llegarían al Nuevo Mundo desde Africa y Europa a finales del Eoceno, cuando el océano Atlántico era más estrecho. Los gaviales vivieron en Europa hasta hace unos 20 millones de años.



PROTOSUQUIOS DE EXTREMIDADES LARGAS, tal como aparecen en un ambiente reconstruido del Triásico superior. Se trata de la especie acorazada *Orthosuchus stormbergi*, procedente de yacimientos fósiles de esta época en Lesotho, en Africa austral, descrita por Diane S. Nash. De 60 centímetros de longitud, debió ser un carnívoro terrestre eficaz pero por aquel entonces se hallaba en fase de adaptación a una vida acuática.

pero en conjunto los crocodilios del sur parecen ser más primitivos. El récord de tamaño entre los Folidosáuridos lo ostenta un enorme ejemplar africano. P. Taquet, del Museo Nacional de Historia Natural de París, ha desenterrado cráneos completos y esqueletos parciales de la especie *Sarcosuchus imperator*, en una notable localidad fósil del Cretácico inferior en la parte sahariana de la República del Níger. Este monstruo parecido a un gavial tenía en total 11 metros de longitud, y su estrecho cráneo medía 1,8 metros. Vivía verosímilmente en los ríos de esta región que ahora es árida.

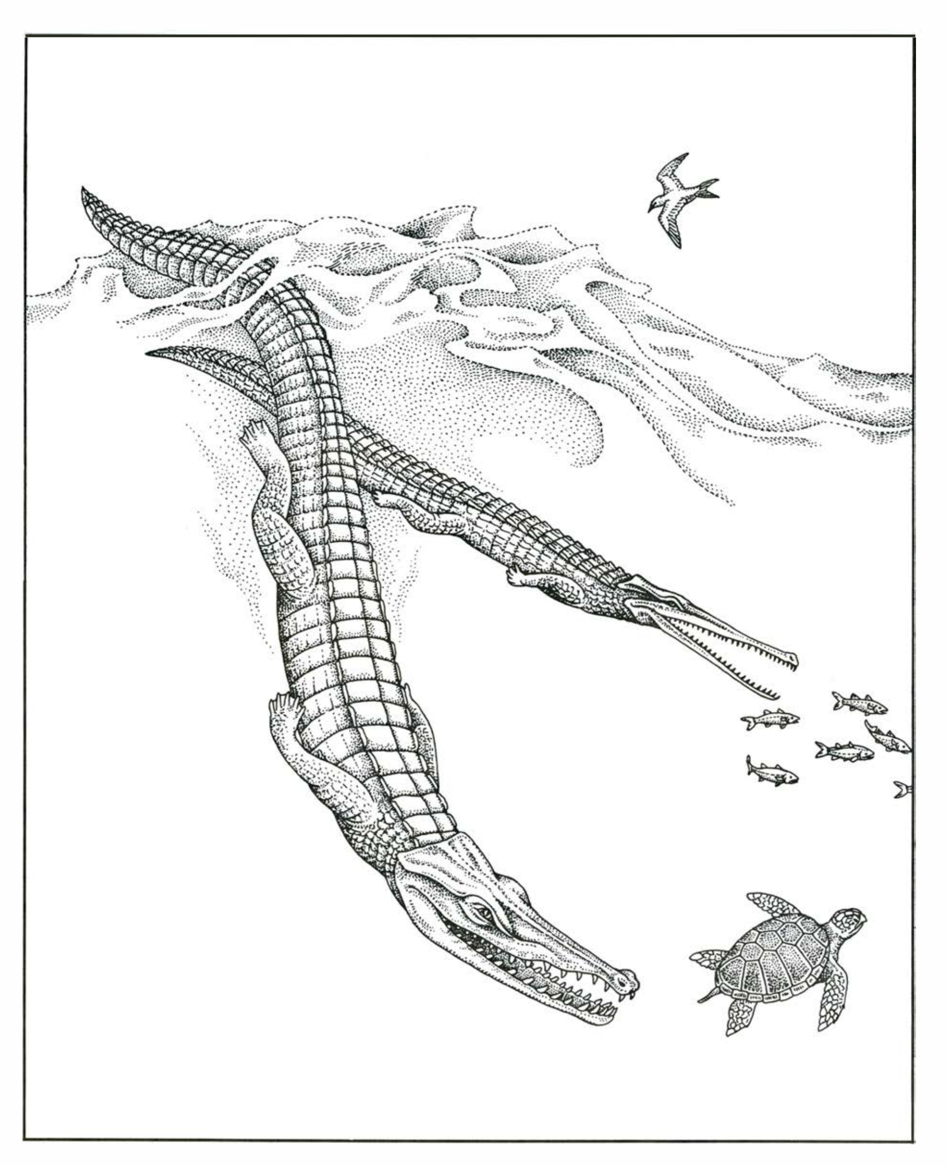
En formaciones del Cretácico inferior de Africa y de Sudamérica se encuentran los mismos peces dulceacuícolas, lo que representa una de entre varias indicaciones de que, por esta época, los dos continentes no se hallaban completamente separados por el brazo de mar del Atlántico Sur. Los crocodilios del Níger proporcionan un apoyo adicional a esta



MESOSUQUIOS ACUATICOS reconstruidos en un ambiente de alta mar del Jurásico. Pertenecen a la familia Metriorrínquidos, la más especializada de los Crocodilios. Habían perdido la armadura de placas y convirtieron sus extremidades en aletas nadadoras. El extremo posterior de la columna vertebral está fuertemente doblado hacia abajo. Impresiones de su piel, conservadas en yacimientos fósiles del Jurásico superior en Alemania, revelan que su cola sostenía una aleta caudal similar a la de los tiburones. La adaptación de los Metriorrínquidos a un nicho de depredador marino empezó en el Jurásico medio; antes de que se extinguieran en el Cretácico inferior, algunas especies habían llegado a alcanzar de cuatro a cinco metros de longitud. Ningún otro crocodilio se adaptó de manera tan perfecta al nicho que dejaron vacío. Vivían en aguas europeas y sudamericanas.

conclusión. Como ejemplo, la colección de fósiles de Taquet procedente del Níger comprende un mesosuquio muy pequeño que pertenece al género *Araripesuchus*; el primer miembro del género que se describió fue desenterrado hace unas dos décadas de una formación del Cretácico inferior del Brasil nororiental. Un segundo ejemplo lo proporcionan los restos de un crocodilio muy grande y de largo hocico del Cretácico inferior de la cuenca de Bahía, que originalmente se clasificó como un goniofólido. Examiné recientemente este ejemplar brasileño y, cuando Taquet y yo comparamos nuestro datos, resultó evidente que pertenecía al género nigeriano *Sarcosuchus*.

Hacia mediados del Cretácico, Africa había quedado completamente separada de Sudamérica, y los crocodilios de los dos continentes divergieron cada vez más, aunque conservaron algunas semejanzas. Al mismo tiempo, pruebas procedentes de ambos continentes demuestran que la evolución de los crocodilios continuó bajo la influencia de las presiones ecológicas durante el resto del período. Tanto en Africa como en Sudamérica surgieron entonces pequeños mesosuquios de hocico corto. Sus restos se encuentran en yacimientos fósiles del Cretácico superior, los Libicosúquidos en el Viejo Mundo y los Notosúquidos en el Nuevo. Con sus mandíbulas cortas, estos animales recuerdan en cierto modo menos a los crocodilios normales que a los mamíferos carnívoros o a los reptiles mamiferiformes. Casi con seguridad ocupaban un nicho de depredador terrestre.

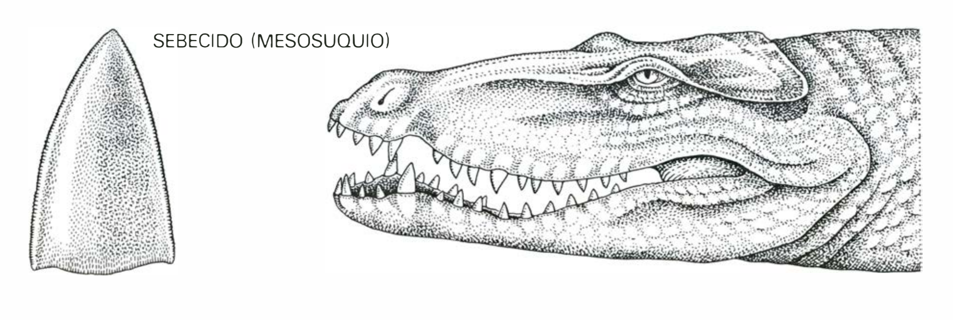


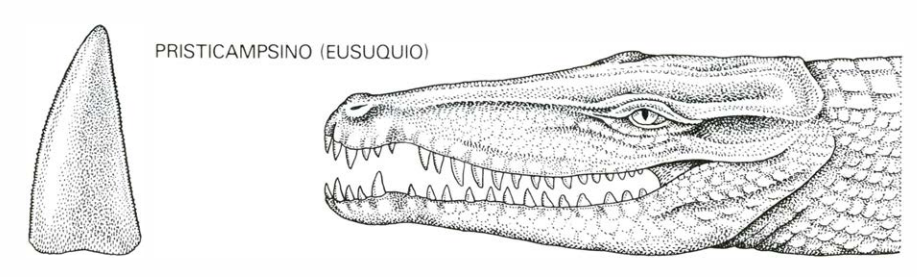
LA DIVERSIFICACION CONTINUA en el Cenozoico afectó a una familia de mesosuquios que apareció en Africa a finales del Cretácico. Se trata de los Dirosáuridos, que habitaban en aguas costeras someras a principios del Terciario. Sus fósiles, abundantes en los yacimientos de fosfatos del sur de Tunicia, muestran que una especie, *Dyrosaurus phosphaticus*, tenía un hocico muy largo y estrecho y dientes puntiagudos adecuados para capturar peces (*derecha*). Su primo, *Phosphatosaurus gavialoides*, tenía un hocico más corto y robusto (*centro*). Sus mandíbulas portaban dientes posteriores grandes y romos, adaptados a la trituración de presas de caparazón duro; esta especie pudo haberse alimentado de tortugas marinas. Las dos especies eran grandes: la piscívora alcanzaba seis metros de longitud, y la que comía tortugas nueve metros. Se han encontrado restos en yacimientos distantes.

No está claro todavía cuál sea el origen de estos crocodilios notables. Su parecido con los Protosuquios primitivos permite sospechar que fueran descendientes directos de estos primeros crocodilios, por vía de un antepasado común que se parecía al mesosuquio primitivo *Araripesuchus*.

El final del Cretácico, hace unos 65 millones de años, estuvo marcado por extinciones en masa de animales a escala mundial. Los Dinosaurios, los reptiles voladores y los grandes reptiles marinos (los Plesiosaurios y los Mosasaurios), todos ellos se extinguieron. De los reptiles arcosaurios sólo sobrevivieron los crocodilios, aparentemente menos afectados por lo que fuera que resultó fatal para muchos otros animales. En esta época se extinguieron algunos crocodilios, como los Goniofólidos, pero probablemente fueron víctimas de la competencia con los Eusuquios, más avanzados, y no de la misteriosa "gran mortandad". (Dos de las tres familias de eusuquios actuales, los Aligatóridos y los Crocodilios, habían aparecido por aquel entonces en Laurasia, en especial en los yacimientos de aguas continentales de 70 millones de años de edad, en Canadá y en los Estados Unidos occidentales.)

La causa de las extinciones en masa ha sido objeto de mucha especulación, y no puede decirse que ninguna sola explicación haya conseguido la aceptación general. El registro fósil de los Crocodilios no revela qué es lo que ocurrió realmente, pero contiene alguna información negativa interesante. Por ejemplo, se ha sugerido que la extinción en masa obedeció a un enfriamiento del clima. Efectivamente, hay algunas pruebas de un empeoramiento del clima hacia finales del Cretácico. Sin embargo, la supervivencia, sin solución de continuidad, de los Crocodilios indica que las temperaturas mundiales no pudieron haber descendido mucho, en particular debido a que algunos de los supervivientes habitaban entonces en latitudes altas. Los crocodilios de hoy en día se hallan confinados a los climas cálidos. Incluso los aligátores, que han desarrollado mecanismos de hibernación, mueren si se hallan sometidos a un período frío prolongado. Un descenso brusco y generalizado de las temperaturas mundiales a fi-





LA EVOLUCION CONVERGENTE produjo dientes planos y aserrados, parecidos a los de los Dinosaurios, en dos estirpes distintas de Crocodilios. En Sudamérica, esta dentición peculiar surgió entre los Sebécidos (*arriba*), un linaje de depredadores terrestres mesosuquios que prosperaron hasta que los mamíferos carnívoros alcanzaron el continente a finales del Terciario. En los continentes septentrionales surgió una dentición carnívora de forma casi idéntica en un grupo de eusuquios depredadores, los Pristicampsinos (*abajo*), que se extinguieron en la parte norte de su distribución durante el Eoceno.

nales del Cretácico habría añadido sin duda los crocodilios a la lista de bajas de los reptiles.

La historia de los Crocodilios ha llegado ahora hasta el Cenozoico. Esta era se divide de modo convencional en un período de tiempo relativamente largo, el Terciario, y un período posterior breve, el Cuaternario. (La terminología procede de los primeros tiempos de la geología histórica, cuando el resto de la historia de la Tierra se encuadraba en las eras Primaria y Secundaria.) Los crocodilios de principios del Terciario continuaron en cierto modo reflejando la dicotomía norte-sur que había surgido durante el Mesozoico. Los Eusuquios estaban bien establecidos en Norteamérica, Europa y Asia; a las poblaciones existentes de antepasados de los cocodrilos y de antepasados de los aligátores y caimanes, se añadieron ahora animales gavialiformes dé hocico largo que vivían en las aguas costeras.

Por esta época, en los continentes australes generalmente aislados un grupo tardío de mesosuquios gozaba de un éxito evolutivo notable. Se trataba de los Dirosáurios, animales grandes que medían hasta nueve metros en total, con mandíbulas largas y estrechas y, en la mayoría de miembros de la familia, dientes largos y puntiagudos que los señalan como adaptados al mismo nicho de piscívoros que había estado ocupado por una sucesión de crocodilios gavialiformes a todo lo largo de la historia evolutiva del orden. Sin embargo, un dirosáurido constituye una variación sobre este tema común. El género *Phosphatosaurus* se encuentra en yacimientos fósiles africanos de principios del Terciario. Sus mandíbulas eran más robustas que las de sus primos piscívoros y sus dientes eran muy fuertes, en particular los posteriores, que poseían puntas redondeadas y romas. Parece que *Phosphatosaurus* hacía presa en animales de caparazón duro que tenían que ser triturados, posiblemente tortugas marinas.

Se han encontrado restos de dirosáuridos en Brasil y en Nueva Jersey, así como en Africa; se conocen asimismo de Asia, donde se encuentran en sedimentos terciarios en Paquistán y Birmania. De hecho, el último dirosáurido fósil que se conoce es una única vértebra que tiene aproximadamente unos 40 millones de años de edad, y que se ha encontrado en un yacimiento fósil del Eoceno tardío en Birmania. Para entonces, la familia había desaparecido ya de Africa, debido presumiblemente a la competencia con los eusuquios septentrionales de hocico largo, adaptados al mismo modo de vida gavialiforme, que se hicieron cada vez más numerosos en Africa durante el Eoceno.

Entre los crocodilios actuales, los gaviales se hallan hoy en día confinados a los grandes ríos del Paquistán, la India y Birmania. Todos pertenecen a la misma especie, *Gavialis gangeticus*. La presencia de gaviales eusuquios ancestrales en Africa durante el Eoceno ayuda a explicar la distribución histórica de los gaviales desde la India hasta Sudamérica (a través de Africa y del Atlántico, que a principios del Terciario era un océano más estrecho). La población africana, que aparentemente había sustituido a los dirosáuridos gavialiformes en el Eoceno, parece haberse visto a su vez presionada por un crocodilio peculiar, con mandíbulas largas y estrechas, que se desarrolló en Africa más avanzado el Terciario y sobrevivió hasta el Cuaternario. Se trata de *Euthecodon*. Su competencia por el mismo nicho ecológico puede explicar la desaparición del gavial de Africa. Por razones que todavía no se conocen, los gaviales americanos y europeos también se extinguieron durante el Cenozoico, de manera que los únicos gaviales supervivientes son los que se refugiaron en los sistemas fluviales del subcontinente indio.

Todos los crocodilios actuales son acuáticos, pero durante el Terciario se hizo prominente en Sudamérica aún otro grupo de crocodilios terrestres. Surgieron evidentemente de un grupo de mesosuquios peculiares, los Baurusúquidos, que aparecieron por primera vez a finales del Cretácico. Al empezar el Terciario habían aparecido sus descendientes, los Sebécidos. Eran animales grandes, con un hocico alto y estrecho. Su rasgo más notable era la dentición: poseían dientes largos y aplanados lateralmente, con bordes cortantes afilados y aserrados, muy parecidos a los de los dinosaurios carnívoros.

Por la época en la que los Sebécidos medraban, Sudamérica quedó separada de los demás continentes. Cuando los mamíferos carnívoros que se habían extendido mucho en otras partes pudieron penetrar en Sudamérica hacia finales del Terciario, su llegada parece haber sido fatal para los sebécidos depredadores. El efecto fue aún más rápido para otro grupo de crocodilios terrestres de dientes de dinosaurio, los Pristicampsinos, que evolucionaron a partir de eusuquios más normales en algún lugar de la región laurásica a principios del Terciario. Aunque los Pristicampsinos tuvieron éxito al principio, con poblaciones distribuidas desde China a Norteamérica, no sobrevivieron a la competencia de los mamíferos más allá del Eoceno, al menos en estas regiones. Existen algunos indicios de que, junto con los Sebécidos, sobrevivieron hasta entrado el Cuaternario en la aislada Australia.

En los continentes septentrionales, la historia de los crocodilios en el Terciario tardío es sobre todo una historia de ex-

tinción. Los aligátores y los gaviales abundaban todavía en Europa a principios del Mioceno, pero al final de esta época, hace unos cinco millones de años, su número comenzó a escasear. Los últimos cocodrilos de Europa desaparecieron, verosímilmente, durante la época siguiente, el Plioceno. Como ha señalado D. E. Berg, de la Universidad de Maguncia, la causa de la decadencia fue sin duda un cambio en el clima europeo: las condiciones se hicieron cada vez más frías y secas durante el Mioceno y el Plioceno, y finalmente los ambientes adecuados para los crocodilios terminaron por desaparecer.

El mismo cambio climático se registró en Norteamérica. Durante el Mioceno vivían aligátores (entre ellos los antepasados probables de la forma americana actual) en lo que hoy son los estados septentrionales centrales de los Estados Unidos. En la actualidad, el aligátor americano se halla confinado a los estados meridionales.

Sin embargo, en las regiones más cálidas del globo los crocodilios continuaron prosperando. Durante el Terciario tardío en Sudamérica evolucionaron varios eusuquios gigantes. Aunque estas formas de enorme tamaño se han extinguido desde entonces, los crocodilios de la región permanecieron pluralmente representados hasta nuestros días, como lo indican las numerosas especies distintas de cocodrilos y de sus primos los caimanes que allí se encuentran. También en Africa los abundantes restos de crocodilios que se hallan en las formaciones fosilíferas del Great Rift Valley demuestran que la población de estos animales era bastante variada hasta bien entrado el Cuaternario. Además de las formas ahora extinguidas, tanto el cocodrilo del Nilo (*Crocodylus niloticus*) como el cocodrilo de Guinea o de hocico largo, de Africa occidental (*C. cataphractus*), se hallaban ya presentes a principios del Cuaternario.

No hay duda de que la depredación por parte del cocodrilo del Nilo constituía un peligro nada despreciable para el hombre primitivo de Africa. En la actualidad la situación se ha invertido: el hombre está desempeñando el papel principal en la decadencia y extinción de los crocodilios. Prácticamente en todas las partes de su distribución, su caza ha diezmado durante años la población de crocodilios. Ha llegado el momento en el que el hombre ha de elegir entre unos cuantos bolsos y zapatos más y la supervivencia de los últimos reptiles arcosaurios que quedan.

Matemática constructivista

Tal planteamiento se basa en la convicción de que la matemática sólo tiene significación real cuando sus conceptos pueden ser construidos por la mente, cuestión que ha escindido al mundo matemático durante más de un siglo

Allan Calder

Suele opinarse que si los humanos llegaran a encontrarse con otra forma de vida inteligente en algún lugar del universo, ambas civilizaciones tendrían en común una matemática fundamental, que podría servir de vía de comunicación. En efecto, desde tiempos de Platón ha sido creencia general que la matemática existe con independencia del conocimiento humano, y que, por tanto, posee un carácter de verdad absoluta. En tal caso, el trabajo de los matemáticos consistiría en descubrir esa verdad. No todos los matemáticos, sin embargo, han compartido esta creencia en la matemática "dada por Dios". Así, Leopold Kronecker, matemático alemán del siglo pasado, sostenía que tan sólo la operación de contar estaba predeterminada. "Dios creó los números enteros", escribió (traduciendo del alemán), pero "todo lo demás es obra del hombre". Según esta concepción, el trabajo de los matemáticos no es descubrir la matemática, sino inventarla.

La cuestión de si la matemática se descubre o se crea no es ociosa. Según la respondan, los matemáticos tendrán enfoques totalmente divergentes sobre cómo debería desarrollarse el trabajo de investigación. En particular, de la creencia de que la matemática se inventa ha surgido la controvertida teoría conocida como matemática constructivista, en la que se sostiene que, para probar que un objeto matemático existe, es necesario mostrar cómo puede construirse. Aunque la teoría constructivista ha sido vertida en odres nuevos y bajo forma sumamente interesante por Errett Bishop, de la Universidad de California en San Diego, sus orígenes pueden hacerse remontar hasta Immanuel Kant, quien ya en el siglo XVIII sostuvo que la razón última de veracidad en la matemática residía en el hecho de que sus nociones podían ser construidas por la mente humana. Quizá puedan comprenderse mejor las consecuencias que sobre las cuestiones de validez de las demostraciones matemáticas y sobre el significado de la noción de existencia matemática entraña semejante punto de vista, si comenzamos examinando las ideas y métodos de los antiguos matemáticos griegos, que han sido guías del desarrollo de esta ciencia a lo largo de los últimos 2500 años.

Para los griegos, la matemática era, en sustancia, geometría, y la quintaesencia de esta disciplina, los *Elementos* de Euclides, donde se conjuntaban los métodos de la escuela filosófica eleática, la lógica de Aristóteles y el idealismo platónico, naciendo la noción de sistema axiomático. Euclides partía de ciertos objetos básicos no definidos, como puntos, rectas y planos, más cierto número de axiomas, es decir, asertos presumiblemente evidentes por sí mismos, concernientes a dichos objetos, y aplicaba las reglas de inferencia aristotélica para deducir nuevas proposiciones, llamadas teoremas. La presunción subyacente, a saber, que por ser "verdaderos" los axiomas y "lógicas" las reglas de inferencia también los teoremas resultarían verdaderos, era plausible, al ser la geometría de Euclides de naturaleza intuitiva. No obstante la imposibilidad de hallar en el mundo natural entes como los puntos (que no poseen tamaño) y las líneas (que no tienen anchura, aunque sí longitud), casi todos creemos saber qué son tales objetos. Además, los teoremas de Euclides podrían comprobarse empíricamente; o sea, podrían realizarse experimentos con lápiz y papel para verificar, pongamos por caso, que dos triángulos cuyos lados tengan longitudes iguales tendrán también ángulos de amplitudes iguales.

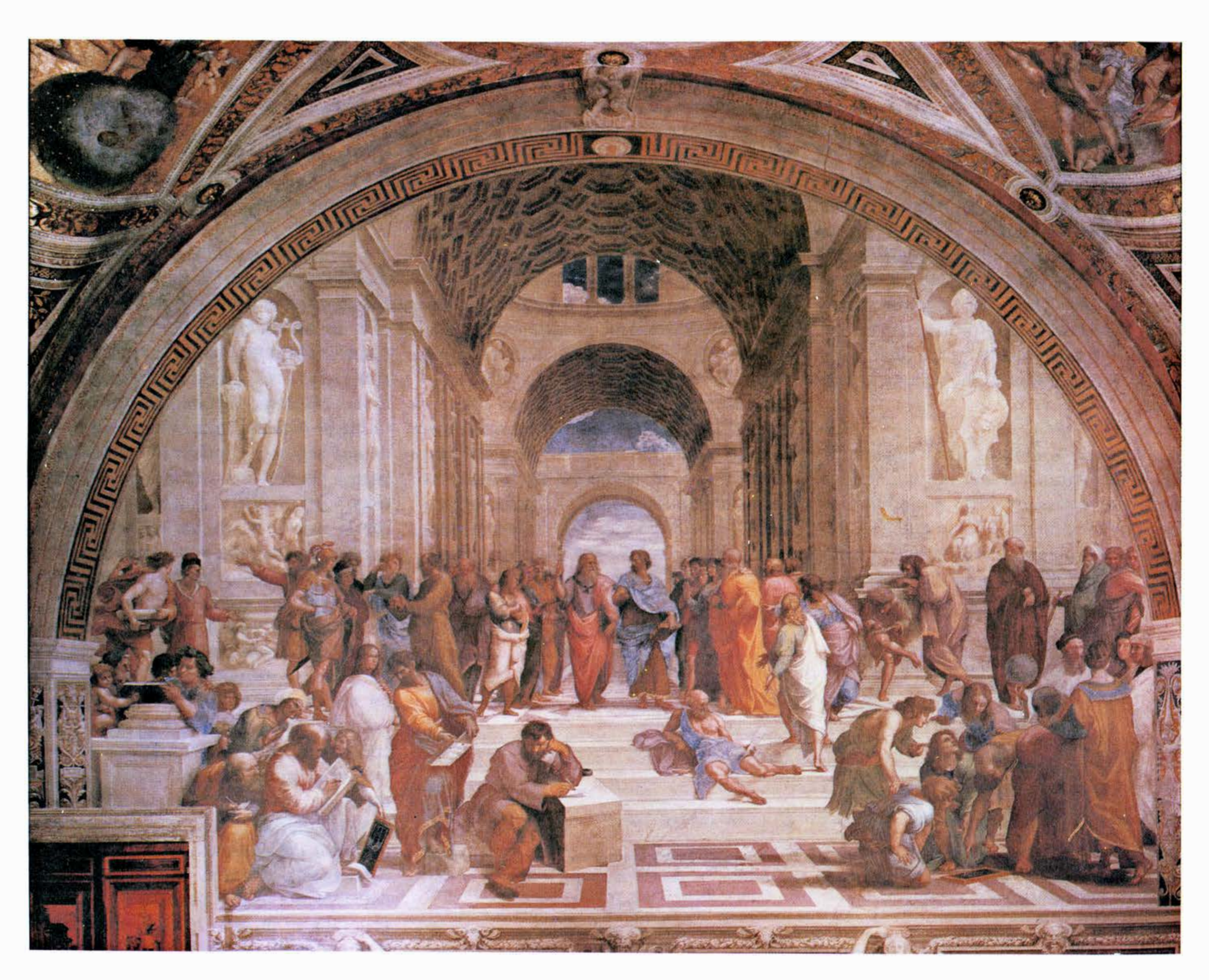
En la era moderna, la matemática ha continuado dependiendo del método axiomático, pero en muchos casos tal método ha sido aplicado a objetos que carecen de las cualidades intuitivas, de la constructibilidad de la geometría de Euclides. Para la escuela constructivista, tales aplicaciones del método han despojado de significación a la matemática. Como resultado, durante los 100 últimos años se han presentado cierto número de propuestas encaminadas a reinstaurar en la matemática el significado intuitivo, la última de las cuales es la matemática constructivista de Bishop. No obstante, los esfuerzos anteriores han contribuido en buena medida a determinar el desarrollo de la matemática, y así, antes de examinar el programa de Bishop, resumiré la historia de esta controversia. Aunque larga y compleja, para comprenderla en sus aspectos fundamentales puede ser suficiente seguir la evolución de una sola idea matemática: la noción de infinito.

Las famosas paradojas de Zenón sobre cuestiones tales como la divisibilidad sucesiva infinita y el continuo indujeron a los antiguos matemáticos griegos a proceder con extrema cautela en lo concerniente al infinito. El objeto fundamental de la matemática griega era el concepto de longitud, y en concreto, Euclides no se refería a rectas ilimitadas, sino tan sólo a segmentos rectilíneos que podían prolongarse hasta alcanzar cualquier longitud deseada. Esta noción de infinito potencial –que para toda línea dada siempre existe otra más larga– resulta de extrapolar razonablemente la experiencia natural. La noción de infinito completo, o infinito actual –a saber, que existen verdaderamente rectas infinitamente largas–, es metafísicamente muy distinta, y exige a la imaginación un atrevido salto al encontrarla por primera vez. No es, por tanto, sorprendente que los matemáticos evitasen manejar infinitos actuales tanto como les fue posible, limitándose a utilizar infinitos potenciales hasta bien entrada la era moderna.

El moderno renacer de la matemática sobrevino cuando, en el siglo XVII, René Descartes y Pierre de Fermat introdujeron las coordenadas cartesianas, transformando así los problemas geométricos

en problemas aritméticos. Al producirse esta aritmetización de la geometría, los objetos fundamentales de la matemática fueron los números, y no las longitudes, lo que preparó el camino que condujo a Isaac Newton y Gottfried Wilhelm Leibniz a desarrollar, cada uno por su parte, el cálculo diferencial: un sistema de reglas para manipular números infinitamente pequeños pero distintos de cero. La noción de magnitud infinitamente pequeña provocó seria desazón (la noción de límite no fue explícitamente formulada hasta mucho más tarde) en una época que atribuía a los objetos matemáticos el mismo grado de realidad que hoy asignaríamos, pongamos por caso, al electrón. De hecho, hubo de transcurrir algún tiempo hasta que la innegable utilidad del cálculo obligó a los matemáticos a aceptarlo, aunque no de buen grado, y con él, los infinitos actuales que implicaba. Durante los 200 años siguientes, los matemáticos se sirvieron con gran éxito de las técnicas del cálculo diferencial, al tiempo que ponían en tela de juicio su base filosófica. Al cabo, este irritante grano de arena incrustado en la ostra matemática produjo a fines del siglo XIX una perla: la teoría de conjuntos.

El desarrollo inicial de la teoría de conjuntos se debe ante todo a Georg Cantor, quien definió los conjuntos como colecciones de objetos reales o abstractos, juntamente con una regla por la cual pudiera decidirse si un objeto dado pertenecía o no al conjunto. Esta idea, tan engañosamente simple, tuvo profundas consecuencias sobre la noción de infinito. Hay conjuntos, como el conjunto de todos los números reales, que, por su misma naturaleza, son infinitos actuales. El estudio de tales objetos llevó a Cantor a la asombrosa conclusión de que, al igual que la cardinalidad de los conjuntos finitos varía de unos a otros, también puede ser variable la car-



EN "LA ESCUELA DE ATENAS", pintado a principios del siglo XVI, Rafael nos muestra los grandes filósofos y matemáticos de la Grecia clásica: Platón (con la mano alzada) y Aristóteles ocupan el centro del cuadro; abajo, a la izquierda, Pitágoras contempla su sistema de proporciones, y Euclides (abajo, derecha) traza una circunferencia en su pizarra. Para los griegos, la matemática era sustancialmente geometría. En el tratado de Euclides, los *Elementos*, se presentó por primera vez la noción del sistema formal axiomático. En este sistema, ciertos enunciados "evidentes *per se*" son aceptados como verdaderos, deduciéndose de ellos nuevas proposiciones, llamadas teoremas, mediante las reglas de inferencia aristotélicas. La naturaleza intuitiva de la geometría euclídea daba fundamento a la presunción de que por ser "evidentes" los axiomas, y "lógicas" las reglas de inferencia, también los teoremas serán verdaderos. El desarrollo de la matemática está todavía guiada por el método axiomático, pero a lo largo de los últimos siglos los objetos estudiados por esta ciencia han ido haciéndose más y más abstractos, al tiempo que menos intuitivos. En consecuencia, algunos matemáticos han puesto en tela de juicio la validez de su disciplina. Muy recientemente, este tipo de ideas ha reverdecido por los trabajos de Errett Bishop, de la Universidad de California en San Diego. Bishop propone reconstruir la matemática sin utilizar nociones no intuitivas, programa que denomina "constructivista". En el sistema de Bishop no es suficiente probar que un objeto matemático ha de existir por razones lógicas, sino que es necesario especificar un procedimiento para construir tal objeto. No es preciso que tal procedimiento sea factible en la práctica, pero sí es necesario que concluya en un número finito de pasos, y que al terminar cada etapa resulte evidente de qué forma proseguir con el paso siguiente.

dinalidad de los conjuntos infinitos. Dicho de otra forma, ciertos conjuntos son más infinitos que otros.

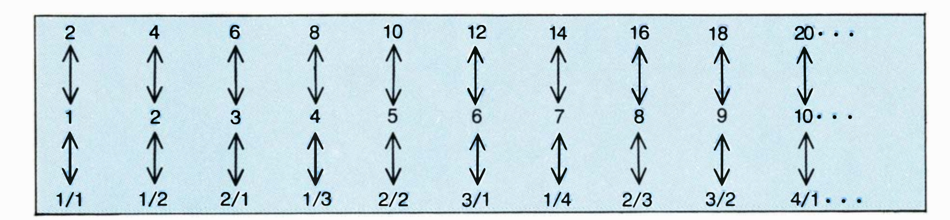
Con mayor precisión, se dice que dos conjuntos tienen igual tamaño, o cardinalidad, cuando los elementos de uno de los conjuntos pueden coordinarse, uno por uno, con los elementos del otro. Y se dice que un conjunto es numerablemente infinito (o contable) cuando tiene la misma cardinalidad que el conjunto 1, 2, 3... de los números naturales. Por ejemplo, los números pares 2, 4, 6... pueden asociarse cada uno con uno de los números naturales, biunívocamente, y forman por consiguiente un conjunto infinito numerable. (Como muestra este ejemplo, un conjunto infinito puede tener la misma cardinalidad que una parte de sí.) Cantor demostró que, para todo conjunto infinito, existe siempre otro de cardinalidad mayor. Por ejemplo, observó que en todo emparejamiento de números naturales con números reales quedarían siempre algunos números reales desemparejados, y por tanto, la cardinalidad de los números reales es estrictamente mayor que la de los números naturales. De hecho, Cantor probó que a partir de los conjuntos numerablemente infinitos puede generarse toda una jerarquía infinita de conjuntos infinitos.

La teoría de conjuntos llegó finalmente a ser cimiento de la matemática moderna, constituyendo los conjuntos –los nuevos objetos fundamentales– poderoso instrumento de abstracción matemática. No obstante, al principio muchos de los más distinguidos matemáticos de entonces encontraron palmariamente absurda la noción de conjunto infinito, como el de los números reales. Incluso el propio Cantor admitió que la idea de considerar conjuntos infinitos como entes individuales le había sido "impuesta por razones lógicas, casi contra mi voluntad". Borrar de esta forma la distinción entre infinito potencial e infinito actual estaba "en contra de tradiciones que habían llegado a ser valiosas". Y a pesar de todo, los conjuntos en sí mismos eran menos inquietantes que las aplicaciones que Cantor vino a darles. Especialmente escandaloso resultó a los matemáticos ortodoxos de la época el procedimiento ideado por Cantor para demostrar que existen infinidad de números trascendentales, también llamados no algebraicos.

Un número real cualquiera x es algebraico cuando verifica alguna ecuación de la forma $a_n x^n + a_{n-1} x^{n-1} + ... + a_1 x + a_0 = 0$, donde $a_1, a_2, ..., a_n$ son números enteros. (Por ejemplo, los números racionales son aquellos números reales expresables en la forma p/q, donde p y q son números enteros y q es distinto de 0. Así pues, todo número racional es algebraico, porque satisface la ecuación $qx - p = 0$ para ciertos valores de p y q.) Cantor probó no sólo que existen números que no satisfacen este requisito, sino que además hay una infinidad de ellos, demostrando primero que el conjunto de números algebraicos es numerablemente infinito. Dicho de otra forma, demostró que esos números pueden ponerse en correspondencia biunívoca con los números naturales. Como hemos visto, el conjunto de los números reales no es numerablemente infinito; por tanto, debe existir números reales "sobrantes", que no serán algebraicos. Supongamos ahora que estos números trascendentales pudieran contarse, o sea, que formasen un conjunto finito o numerablemente infinito. No es difícil demostrar que un conjunto obtenido por reunión de dos conjuntos contables es también contable, numerable. (No hay problema cuando uno de ellos, o ambos, sean finitos. Si ambos conjuntos son infinitos, para probar que la unión es numerable basta poner uno en correspondencia biunívoca con los números impares, y el otro con los números pares.) Así pues, si la hipótesis fuese verdadera, el conjunto de todos los números reales sería contable (por ser reunión de los números algebraicos y los no algebraicos) –conclusión falsa. Por consiguiente, la hipótesis de que los números trascendentales forman colección numerable es falsa. Tienen que ser infinitos y no numerables.

La demostración de Cantor se apoya en el clásico método de *reductio ad absurdum*: para probar que una proposición es verdadera se demuestra que su negación conduce a contradicción. Para comprender la airada reacción que matemáticos como Kronecker opusieron a tal demostración hace falta darse cuenta de que según la tradición matemática no podría haber existencia sin construcción. Con precisión mayor, para establecer la existencia de un objeto matemático se consideraba necesario especificar un procedimiento mediante el cual el objeto pudiera ser construido, al menos en principio. No era preciso que tal procedimiento fuese llevado a la práctica, ni tampoco se requería que pudiese serlo en un tiempo "razonable". Las únicas estipulaciones eran que el procedimiento concluyese en número finito de pasos, y que al final de cada etapa no hubiese duda de cómo proceder a ejecutar el paso siguiente. En su demostración de infinitud del conjunto de números trascendentales, Cantor incumplía escandalosamente estos requisitos, pues no generaba ni un solo ejemplo de número trascendental, y no digamos ya una infinidad. (Resulta interesante destacar que ni siquiera la definición tradicional de existencia era considerada suficientemente estricta por todos. Se sabe de un matemático que invirtió 10 años en construir con sólo regla y compás un polígono regular de 65.537 lados, objeto cuya existencia había sido establecida por Carl Friedrich Gauss.)

La demostración de Cantor fue el primer ejemplo importante de las que se ha



LA TEORIA DE CONJUNTOS INFINITOS, que fue desarrollada por Cantor en la década de 1870, instaura como objetos fundamentales de estudio matemático los conjuntos, papel que inicialmente desempeñaron las longitudes, y más tarde, los números. Cantor hizo ver que ciertos conjuntos infinitos, como los números pares (*fila superior*) y los números racionales (*fila inferior*) son contables (o numerables), en el sentido de que pueden ponerse en correspondencia biunívoca con los números naturales 1, 2, 3, ... (*fila central*), mientras que otros conjuntos infinitos, como el de los números reales, no pueden ser contados, pues son esencialmente mayores que el conjunto de los números naturales. (Los números racionales, o sea, el conjunto de todas las fracciones, no tienen una ordenación natural; aquí los hemos ordenado según la suma de su numerador y denominador.) Cantor ideó una serie de números "transfinitos" para expresar la cardinalidad o número de elementos de estos conjuntos infinitos, y después elaboró una aritmética general, apta tanto para números finitos como transfinitos. Antes de Cantor, la mayoría de los matemáticos se limitaban a la noción de infinito potencial, a saber, que para todo número siempre existe otro mayor. Los conjuntos infinitos obligaron a los matemáticos a enfrentarse con la noción de infinito actual, totalmente distinta. Un infinito actual, o completo, es una colección formada por una infinitud de objetos que pueden ser tomados todos de una vez. Los matemáticos de la escuela constructivista, defendida en este artículo, opinan que los números, y no los conjuntos, son la verdadera base de la matemática. Como la naturaleza parece no contener conjuntos infinitos, los matemáticos no han podido adquirir la comprensión real, intuitiva, en que fundar los juicios concernientes a ellos. Resulta así que muchas de las más fecundas y conocidas técnicas de la matemática moderna tendrían que ser profundamente modificadas para ser válidas en teoría constructivista.

dado en llamar demostraciones de pura existencia. Al no dar la más leve indicación de cómo podría hallarse ni siquiera uno de estos números trascendentales, establecía la existencia de tales números demostrando que sería contradictorio que no existieran. Como antes, la pieza clave es el infinito. Las demostraciones por reducción al absurdo que establezcan la existencia de objetos en conjuntos finitos son perfectamente aceptables por todos los matemáticos; pues en principio se podría exhibir tal objeto inspeccionando uno por uno los miembros del conjunto. Mas no podría hacerse así para, pongamos por caso, el conjunto de números trascendentales, quienes pertenecen al conjunto infinito de los números reales. Por esta razón, muchos matemáticos rechazaron de plano la demostración de Cantor, objetando que una contradicción nunca podrá remplazar a un ejemplo tangible. La polémica estaba lejos de resolverse cuando en 1889 David Hilbert publicó una demostración de pura existencia todavía más escandalosa.

La demostración concernía a un problema propuesto por el matemático alemán Paul Gordon. El problema de Gordon, una de las más destacadas cuestiones abiertas del tiempo, pedía demostrar la existencia de ciertos conjuntos (finitos) de objetos dotados de propiedades especiales, en cierto número de casos. Hilbert demostró por *reductio ad absurdum* que los conjuntos requeridos existían para cada uno de tales casos. Cantor había demostrado la existencia de un conjunto de objetos de los que al menos se disponía de algunos ejemplos; en efecto, ciertos números trascendentales habían sido descubiertos por lo menos 30 años antes. Y ahora Hilbert sostenía haber probado la existencia de objetos que nadie había visto jamás y de los que no se tenía la menor idea para construirlos. "Esto no es matemática", dijo Gordon. "Es teología".

Sin embargo, en manos de Hilbert las demostraciones de pura existencia se convirtieron pronto en potente herramienta, y conforme la teoría de conjuntos fue produciendo valiosos resultados matemáticos, fue remitiendo la violenta objeción a sus métodos. "La teología", llegó a conceder Gordon, "también tiene sus méritos". Tras la muerte de Kronecker, en 1891, la oposición a métodos no constructivistas fue vacilante, hasta encontrar nueva voz en el matemático holandés L. E. J. Brouwer. En opinión de Brouwer, la matemática llegaría a carecer de significado a menos que se efectuase una distinción tajante entre la existencia en sentido real, constructivo, y la existencia "pura". Ambas nociones habían llegado a unificarse al generalizar la lógica clásica –específicamente, la ley de tercero excluido– y aplicarla no sólo a conjuntos finitos sino también a los conjuntos infinitos.

La ley aristotélica del tercero excluido afirma que cada proposición es verdadera o falsa; sin ella es imposible toda demostración por contradicción. Por ejemplo, en la demostración de Cantor de existencia de números trascendentales se prueba que la hipótesis de ser algebraicos todos los números conduce a contradicción; por consiguiente, según la ley del tercero excluido, debe existir números no algebraicos, o sea, trascendentales. La lógica aristotélica forma parte en tan gran medida de la vida ordinaria que puede parecer absurdo poner su validez en tela de juicio. Casi todo el mundo está convencido de que la ley del tercero excluido es correcta, pues la experiencia cotidiana la confirma en todo momento. Por ejemplo, dada una bolsa que contiene cierto número de bolas de colores, nadie podría objetar nada si se le dijese que en la bolsa o bien hay una bola roja o bien no la hay. Dicho de otra forma, no es necesario mirar en la bolsa para comprobar que uno de los dos asertos es verdadero. Y para decidir cuál de ellos lo es, basta ir observando el contenido de la bolsa; tras un tiempo finito la cuestión quedará aclarada.

La situación es totalmente distinta al aplicar la ley del tercero excluido a conjuntos infinitos. Imaginemos una bolsa que contuviera una infinidad de bolas de colores. En este caso, tanto si las bolas van siendo observadas de una en una, como si lo son de mil en mil, nunca será posible examinarlas todas. Por consiguiente, aunque sí podría ocurrir que en algún momento del proceso de inspección se pudiera afirmar que la bolsa sí contiene una bola roja, nunca será posible decir que la bolsa no contiene bolas rojas. Tal es la dificultad en que nos coloca la noción de infinito: no importa lo mucho que se haya inspeccionado, siempre habrá más por ver. La única forma de saber que la bolsa no contiene ninguna bola roja sería "ver" todas las infinitas bolas de una vez.

Puesto que no se han encontrado en la naturaleza infinitos actuales, los matemáticos no disponen de conocimiento real e intuitivo acerca de ellos. Así pues,

Proposición: Existe una infinidad de números primos.

Demostración de Euclides

1. Sea $p_1, p_2, \ldots, p_m$ cualquier conjunto finito de números primos.

2. Tomemos el número $p_1 \times p_2 \times \ldots p_m + 1$ que no puede ser dividido exactamente por ninguno de los primos $p_1, p_2 \ldots, p_m$. O bien este número es primo, o bien es divisible por algún otro número primo p.

3. En cualquiera de estos casos, tiene que existir un primo que no pertenece al conjunto $p_1, p_2, \ldots, p_m$. Por tanto, el conjunto de los números primos es infinito.

Demostración moderna

1. Supongamos que la proposición sea falsa, o sea, que la colección de números primos sea finita. Supongámosla formada por $p_1, p_2, \ldots, p_m$.

2. Tomemos el número $p_1 \times p_2 \times \ldots p_m + 1$, que no es divisible exactamente por ninguno de los números $p_1, p_2, \ldots, p_m$. O bien este número es primo, o bien es divisible por algún otro número primo p.

3. En cualquiera de estos casos tiene que existir un número primo no perteneciente al conjunto $p_1, p_2, \ldots, p_m$. Este resultado contradice la hipótesis de ser finito el conjunto de números primos; luego la hipótesis ha de ser falsa. Así que (por el principio de tercero excluido) el conjunto de números primos es infinito.

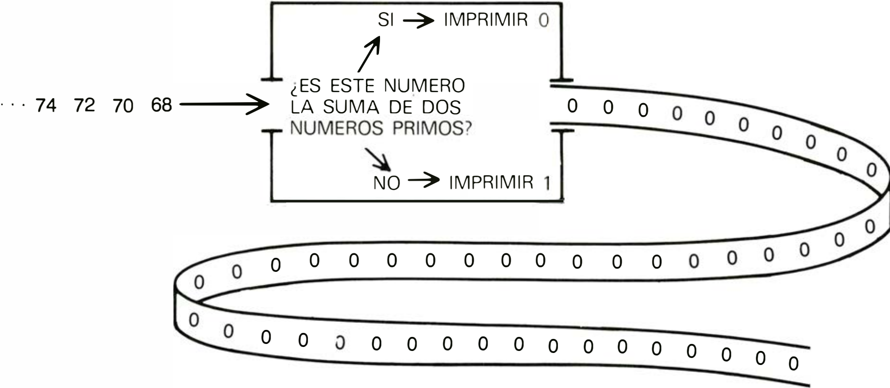
DEMOSTRACIONES CONSTRUCTIVISTAS Y NO CONSTRUCTIVISTAS: se distinguen por la forma en que utilizan el infinito, como podemos apreciar comparando la primitiva demostración sobre la existencia de infinitos números primos, dada por Euclides, con la demostración moderna habitual de esta misma proposición. (Un número primo es un número entero positivo que únicamente es divisible por 1 y por sí mismo.) La demostración de Euclides (*izquierda*) es constructivista, y pone de manifiesto que para todo grupo finito de números primos existe siempre un grupo mayor; la moderna, no constructivista (*derecha*), prueba la existencia de una infinidad actual de números primos. La demostración no constructivista se apoya en la técnica de reducción al absurdo, a saber, para probar que una proposición es verdadera se demuestra que al negarla resulta una contradicción. Las demostraciones por contradicción descansan en un principio aristotélico, la ley del tercero excluido, que afirma que dada una proposición cualquiera, o bien es verdadera, o bien es falsa. Los matemáticos constructivistas impugnan la aplicación de esta ley en el caso de conjuntos infinitos, alegando que podría ser imposible llegar a decidir cuál de las alternativas es la verdadera ***(véase la ilustración de la página siguiente)*****.**

para aplicar a conjuntos infinitos la ley de tercero excluido forzosamente tienen que extrapolar su experiencia con conjuntos finitos. Tal extrapolación resulta particularmente difícil de justificar si tenemos en cuenta que los griegos que idearon la ley del tercero excluido contemplaban con la mayor suspicacia los infinitos actuales. En efecto, como concluía Brouwer, los resultados de extrapolar así, a partir de la experiencia de casos finitos, no tienen por qué ser "constructivamente" verdaderos; es decir, que es posible demostrar la existencia de objetos que no pueden ser construidos [*véase la ilustración de esta misma página*]. Brouwer aducía que, de atribuir a tales objetos ideales la misma validez que a los objetos reales –esto es, finitamente constructibles–, la matemática perdería su certidumbre. Los entes ideales serían utilizados para crear nuevos entes ideales, y pronto no habría forma de saber si el tema estudiado poseería algún significado real.

La crítica de Brouwer se produjo en un momento en que la teoría de conjuntos era particularmente vulnerable. Después de 1900, la reorganización de la matemática sobre los fundamentos de la teoría de conjuntos había dado, entre otros frutos, varias paradojas no elucidadas. Por ejemplo, ya el propio Cantor observó que, como para todo conjunto existe otro de cardinalidad mayor, tendría que existir un conjunto V de cardinalidad mayor que la del conjunto de todos los conjuntos, U, en el cual está contenido todo. Por otra parte, como U contiene todo, los elementos de V han de ser elementos de U y, por tanto, la cardinalidad de U tiene que ser mayor o igual que la de V. Paradojas como ésta indicaban que los fundamentos teórico-conjuntistas de la matemática debían contener defectos; y Brouwer afirmaba que los defectos obedecían precisamente a la introducción de objetos ideales.

Brouwer estaba convencido de que la crisis provocada por estas paradojas tendría que resolverse expurgando la matemática de razonamientos y objetos ideales, es decir, reconstruyendo totalmente la matemática según procedimientos estrictamente constructivistas. Dado que este programa, llamado intuicionismo, hubiera dado al traste con la mayor parte de la matemática contemporánea, incluso matemáticos simpatizantes con la crítica de Brouwer se resistían a ponerlo en práctica. En cambio, fueron mucho más receptivos a un plan ideado por Hilbert, no para depurar la matemática desde su interior, sino para justificarla desde afuera. De esta forma, los matemáticos no se verían –así lo expresó Hilbert– expulsados del paraíso conjuntista que Cantor creó para ellos.

En opinión de Hilbert, aunque los objetos ideales pudieran carecer de significado por sí mismos, todavía podrían deducirse de ellos objetos y teoremas con significado. Desde su punto de vista, podría restaurarse la certidumbre de la matemática considerándola como sistema puramente formal: un juego cuyo interés reside, no en los objetos que se manejan, sino en sus reglas y en las relaciones que ellas crean entre los objetos. Quedará demostrado que la matemática tiene significado tan pronto como se demuestre que el sistema sea coherente, esto es, esté libre de contradicciones, sostenía Hilbert.



CONJETURA DE GOLDBACH, que enuncia que todo número entero par es suma de dos números primos. Aunque esta famosa proposición nunca ha sido demostrada, existe un algoritmo eficiente para determinar si un número par dado es, o no, suma de dos números primos. Así pues, sería tarea sencilla programar un ordenador que fuera comprobando uno por uno los números pares, e imprimiera un 0 si el número resultase suma de dos primos, y un 1 en caso contrario, como muestra la ilustración. Por ser infinito el conjunto de los números pares, la sucesión de ceros (y quizá de unos) generada por el ordenador sería también infinita. ¿Es lícito aplicar a esta sucesión infinita el principio aristotélico del tercero excluido, y afirmar que o bien la sucesión contiene un 1, o bien no lo contiene? La dificultad reside en que tal vez nunca sea posible decidir cuál de las alternativas es correcta. Si resultase impreso un 1, podríamos afirmar concluyentemente que la sucesión sí contiene un 1 (y que la conjetura es falsa); pero por mucho tiempo que esté funcionando el ordenador, mientras sólo imprima ceros no se podrá nunca estar seguro de que más adelante no saldra un 1 (ni se podrá decir que la conjetura es verdadera). En matemática constructivista únicamente se puede afirmar que la sucesión, o bien contiene un 1 o bien no lo contiene, cuando pueda especificarse explícitamente un procedimiento (finito) para decidir cuál de las alternativas es verdadera. Los matemáticos constructivistas están convencidos de que todo problema matemático puede ser reducido a un problema numérico análogo a la conjetura de Goldbach, y que toda cuestión aritmética puede ser reducida al problema de decidir si en una sucesión compuesta por ceros y unos se presenta o no un 1. Por consiguiente, si fuese posible demostrar constructivistamente que toda sucesión de ceros y unos, o bien contiene un 1 o bien no lo contiene, todos los problemas matemáticos podrían ser resueltos, pues tal demostración tendría que contener un procedimiento finito para determinar qué alternativa es verdadera. No obstante, la mayoría de los matemáticos opinan que tal procedimiento no puede existir: que no se hallará una demostración constructivista que lo abarque todo.

El programa de Hilbert, generalmente denominado formalismo, requería ante todo representar la matemática como un sistema formal axiomático. Se tenía confianza en que transformando los enunciados matemáticos en cadenas de símbolos desprovistos de sentido, y combinados de acuerdo con las reglas de la lógica, todo principio de razonamiento no manifiesto o implícito que hubiera podido provocar paradojas quedaría al descubierto. Todavía más importante, una vez que la matemática estuviese formalizada por completo se podría acometer la demostración de que es coherente. A tal objeto, Hilbert introdujo una nueva teoría, llamada teoría de demostración, o metamatemática, en la que podían formularse enunciados provistos de significado concernientes a las configuraciones y símbolos del sistema axiomático, desprovistos de él. Los razonamientos y símbolos de la metamatemática tenían obligatoriamente que ser de naturaleza totalmente constructivista, de un número finito de pasos. De esta forma no podría ponerse en duda la validez de una demostración metamatemática de la coherencia de la matemática.

Brouwer no aceptó la proposición de que al demostrar la coherencia de la matemática quedaría restaurada su significación. Y dedicó casi toda su vida a mostrar que la matemática podía realizarse constructivistamente. Empero, su programa nunca alcanzó gran apoyo, no siendo la menor de las razones que los esfuerzos matemáticos de los intuicionistas resultaban más desmañados y artificiosos. Y cuando Brouwer intentó depurar la matemática de nociones ideales, al tiempo que se esforzaba en respetar algunas de sus características adquiridas desde antiguo, introdujo ideas que algunos juzgaron tan poco realistas como las que trataba de eliminar.

Irónicamente, no fue la gran fe en la

matemática lo que provocó el rechazo de la crítica intuicionista, sino más la falta de tal fe. Aunque muchos temían que su disciplina estaba efectivamente perdiendo significación, parecía como si la única manera de mantenerla viva fuese aceptar ciertas nociones ideales. Como expuso Hilbert: "Comparada con la inmensa extensión de la matemática moderna, ¿qué podrían significar miserables residuos, los escasos y aislados resultados, inconexos e imcompletos, que los intuicionistas han conseguido?". Para la mayoría de los matemáticos, el formalismo de Hilbert ofrecía un compromiso aceptable y atractivo. Si pudiera demostrarse que la matemática no contiene contradicciones, al menos estaríamos seguros de que si dos personas (o una sola en dos ocasiones diferentes) realizasen el mismo juego matemático no se presentarían con dos soluciones distintas y correctas.

Y entonces, en 1931, Kurt Gödel demostró que el esquema formalista de Hilbert estaba condenado al fracaso. Gödel demostró que no se puede probar la coherencia de ningún sistema formal axiomático lo bastante rico como para contener la aritmética mediante razonamientos metamatemáticos. Lo más que se puede esperar es saber que tal sistema es incoherente. Sin embargo, en tal fecha la teoría de conjuntos y el formalismo estaban tan integrados en la matemática que era imposible rechazarlos sólo porque la meta fundamental del formalismo se probase inaccesible.

A lo largo del último medio siglo, el formalismo ha llegado a constituir parte integral del pensamiento matemático, mientras que el intuicionismo ha venido a ser considerado como pintoresca curiosidad matemática. Hoy por hoy, la mayoría de los matemáticos prefieren la pulcra elegancia de las demostraciones de pura existencia, aún cuando dispongan de demostraciones constructivistas de idénticos resultados; rara vez se hace cuestión de la certidumbre de tal matemática. Todavía hay cierto debate sobre unos cuantos aspectos de la teoría (en particular, el empleo del axioma de elección, que se examina más adelante), pero en su mayor parte la polémica relativa a los métodos conjuntista de Hilbert y Cantor se ha calmado. Quizá por este motivo llame más poderosamente la atención que se está produciendo ahora un reverdecimiento de la matemática constructivista.

Este renovado interés data de la publicación, en 1967, del libro de Errett Bishop, *The Foundations of Constructive Mathematics*. Si bien los esfuerzos de los más polémicos precursores de Bishop fueron juzgados desmañados y artificiosos, el trabajo de Bishop pone de relieve que los resultados obtenidos por métodos constructivistas pueden ser tan bellos y útiles como los alcanzados por métodos formalistas. Brouwer quiso reforzar los cimientos de la certidumbre matemática, y comenzó con un sistema intuitivo de axiomas para la aritmética, ascendiendo a partir de ahí. Pero el plan de Bishop para "constructivizar" la matemática tiene distinta orientación. Su preocupación no es fundamentar la matemática, sino desarrollarla constructivísticamente. En consecuencia, parte ya de los números racionales y la aritmética ordinaria, proponiéndose continuar "realísticamente". En la concepción de Bishop, la certidumbre puede alcanzarse evitando toda noción ideal e inspeccionando sistemáticamente el significado intuitivo de los objetos y teoremas engendrados. Sus resultados muestran que podemos desarrollar gran cantidad de valiosa matemática con efectividad y eficacia mediante estos métodos directos.

Por ejemplo, en lugar de rechazar la teoría de conjuntos cantoriana, como hizo Brouwer, Bishop modifica la teoría con objeto de dotarla de validez constructivista. Concretamente, insiste en que, para definir conjuntos, no basta simplemente con dar una regla por la que se decida la pertenencia al conjunto. Bishop exige también que se den procedimientos para construir efectivamente miembros del conjunto, y para demostrar que dos elementos son distintos. Interesante resultado de estas modificaciones es que, con ellas, uno de los más debatidos postulados de la teoría de conjuntos cantoriana, el axioma de elección, resulta totalmente aceptable.

El axioma de elección afirma que, para toda colección no vacía de conjuntos no vacíos (la colección tiene al menos un conjunto, y cada uno de sus conjuntos, al menos un elemento), siempre es posible elegir un representante en cada conjunto. Dicho con otras palabras, cabe dar una regla de selección tal que dos personas que examinasen la colección e hicieran su elección de acuerdo con la regla, producirían el mismo conjunto de representantes. Para comprender por qué este axioma, de empleo universal en matemática, se encuentra en entredicho, tomaremos el siguiente ejemplo, debido a Bertrand Russell.

Para cualquier colección infinita de pares de zapatos, la regla "Elíjase el zapato del pie izquierdo" satisfaría el axioma de elección, pero, ¿qué regla serviría para una colección infinita de pares de calcetines? Si solamente hubiese un número finito de pares de calcetines, se podría revisar la colección e ir marcando un calcetín de cada par, con lo que la regla "Elíjase el calcetín marcado" sería suficiente. No obstante, como existen infinitos pares, no hay forma de marcar todos los calcetines antes de empezar y, por tanto, parece no existir regla de elección. Además, el axioma de elección puede también presentar dificultades cuando se aplica a colecciones formadas por un número finito de conjuntos, e incluso, de un conjunto tan sólo. Por ejemplo, una de las más famosas conjeturas no demostradas de la teoría de números es el teorema magno de Fermat. Se podría definir que un conjunto está formado por 0 y 1 si el teorema de Fermat es verdadero, y por 2 y 3 si es falso. En este ejemplo, la regla "Elíjase el menor de los números del conjunto" determinaría unívocamente el representante del conjunto, pero hasta que el teorema sea probado, no hay forma de especificar el número concreto que representa al conjunto.

En la revisión constructivista efectuada por Bishop de la teoría de conjuntos, sin embargo, el axioma de elección no resulta impugnable, porque se estipula que, para declarar un conjunto como no vacío, hay que demostrar que el método de construcción especificado al definirlo genera efectivamente un elemento del conjunto. Así pues, en la teoría constructivista puede suponerse que, en cada conjunto no vacío, existe un elemento distinguido, a saber, el construido al demostrar su no vacuidad. El axioma de elección queda satisfecho sin más que tomar el elemento distinguido de cada conjunto de la colección.

Para darnos cuenta de algunas de las ramificaciones de la teoría de conjuntos constructivista de Bishop, fijémonos en el conjunto de números reales. Para definir constructivamente este conjunto es necesario especificar un método de construcción de los reales a partir de los números racionales, intuitivamente válidos. O sea, es necesario construir números irracionales, como $\sqrt{2}$ y e a partir de los números racionales. Podría suponerse que tales números estarían construidos tan pronto como se dispusiera de un procedimiento para calcular su desarrollo decimal con tantas cifras como se desee. Sin embargo, esta definición de "construcción" no es satisfactoria, porque el conjunto de números obtenido por este procedimiento no es cerrado frente a las operaciones fundamentales de la aritmética, como forzosamente han de serlo los números reales. Es decir, con esta definición, la suma de dos nú-

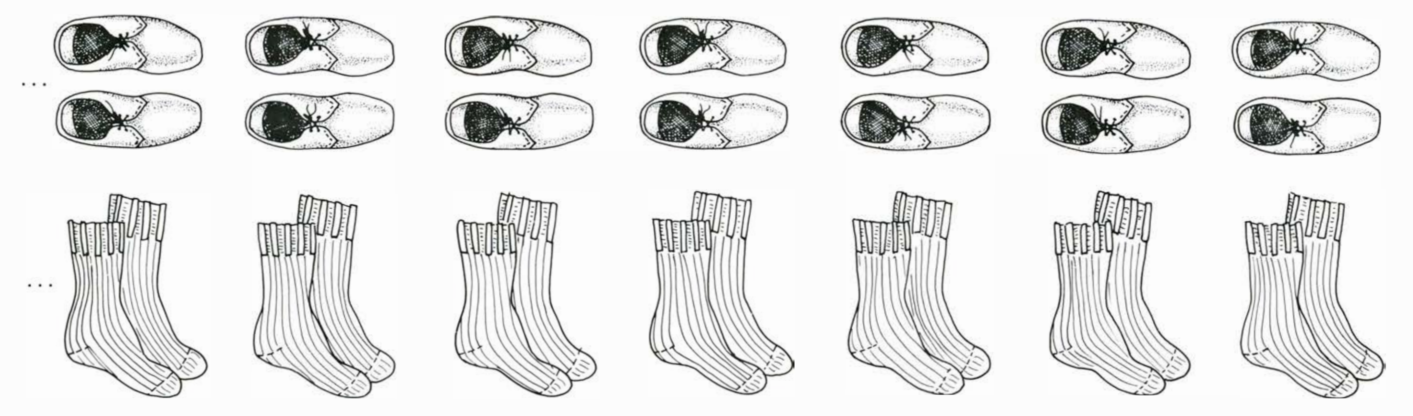
meros reales no es necesariamente un número real, el producto de dos números reales no es necesariamente un número real, y así sucesivamente. El siguiente ejemplo muestra cómo podrían construirse por este procedimiento dos números a y b, sin que pueda construirse su suma.

Se ignora si en el desarrollo decimal del número irracional π figura o no la secuencia de cifras 0123456789. Además, admitiendo que sí se presente esta sucesión, pongamos por caso, por primera vez a partir del "número crítico" k, entonces se ignora si k es par o impar. Empero, hay un procedimiento puramente mecánico para determinar si un número entero n es o no es número crítico de π: basta con calcular π hasta la $n+9$-ésima cifra decimal. Por consiguiente, podemos definir a como el número decimal cuyo desarrollo comienza 0,333... y prosigue indefinidamente de esta forma a menos que algún número impar n sea número crítico de π; en tal caso se coloca un 4 en la n-ésima cifra decimal y en todas las situadas a su derecha. Análogamente, b puede ser definido como el número que comienza por 0,666... y continúa de esta forma a menos que algún número par n sea número crítico de π; en tal caso se colocaría un 5 en la n-ésima cifra decimal y en todas las posteriores.

Notemos que si π no tiene número crítico, entonces a es igual a $1/3$, b es igual a $2/3$ y $a+b$ es igual a 1. Por otra parte, si el número crítico de π existe y es par, entonces a es mayor que $1/3$, b es igual a $2/3$ y $a+b$ es mayor que 1; mientras que si el número crítico de π existe y es impar, a es igual a $1/3$, b es menor de $2/3$ y $a+b$ es menor que 1. Intentemos ahora empezar a escribir el desarrollo decimal de $a+b$. Si la suma comienza por 1, entonces no podrá ser menor que 1, y el número crítico de π, si existe, ha de ser par; si la suma comienza por 0,9, no podrá ser mayor que 1, y el número crítico de π, si existe, ha de ser impar. Puesto que nadie tiene la menor idea de cómo demostrar que el número crítico de π, caso de que exista, sea ya par, ya impar, no es posible escribir ni tan siquiera la primera cifra decimal del desarrollo de $a+b$. Tal como este ejemplo pone de manifiesto, pensar en definir números reales explicando cómo calcular su desarrollo decimal resulta ser idea excesivamente ingenua, tanto si nos ajustamos a criterios constructivistas como si no.

¿De qué forma podremos entonces construir el sistema de los números reales? Tanto desde el punto de vista constructivista como desde el formalista, un número real queda determinado cuando es posible aproximarlo tanto como se desee mediante números racionales. Por ejemplo, el desarrollo decimal de un número x hasta, pongamos por caso, $n+1$ cifras decimales se diferencia de x en menos de $1/10^n$. (Fijémonos en el conocido desarrollo decimal de π, que comienza por 3,141592...; tomando este desarrollo hasta el quinto decimal resulta 3,14159, número que difiere de π en 0,00002..., que es menor que 0,0001, es decir, que $1/10^4$.) Tomando un valor de n suficientemente grande se puede siempre obtener de esta forma un número racional que se encuentre todo lo cercano a x que se desee. Como hemos visto, la idea de utilizar solamente números decimales para representar números reales no es suficientemente fructífera. Por contra, hay otros muchos procedimientos para obtener aproximaciones racionales de números reales. Admitiendo como definido todo número real que pueda aproximarse por algunos de tales procedimientos (y no solamente mediante desarrollos decimales) se consigue un sistema perfectamente suficiente. En la teoría constructivista, lo anterior significa que se admite construido un número real cuando se especifica un procedimiento (finito) de obtención de aproximaciones racionales hasta cualquier grado de precisión dado de antemano. Dicho de otra forma, ha de ser posible obtener en un número finito de pasos un número racional que se diferencie del número real en menos que dicho grado. Por ejemplo, para aproximar el número $a+b$ anterior, cuyo desarrollo decimal no sabíamos escribir, con precisión mejor que $1/10^n$ basta simplemente calcular π hasta $n+10$ cifras decimales. Si este cálculo muestra que el número crítico de π es menor o igual que $n+1$, entonces $a+b$ puede ser calculado exactamente hasta $n+1$ cifras decimales. En caso contrario, a tendrá la forma 0,333..., y b será 0,666... hasta por lo menos la $n+1$-ésima cifra decimal y, por ello, el número racional 1 se diferenciará de $a+b$ en menos de $1/10^n$.

Con mayor generalidad, los números

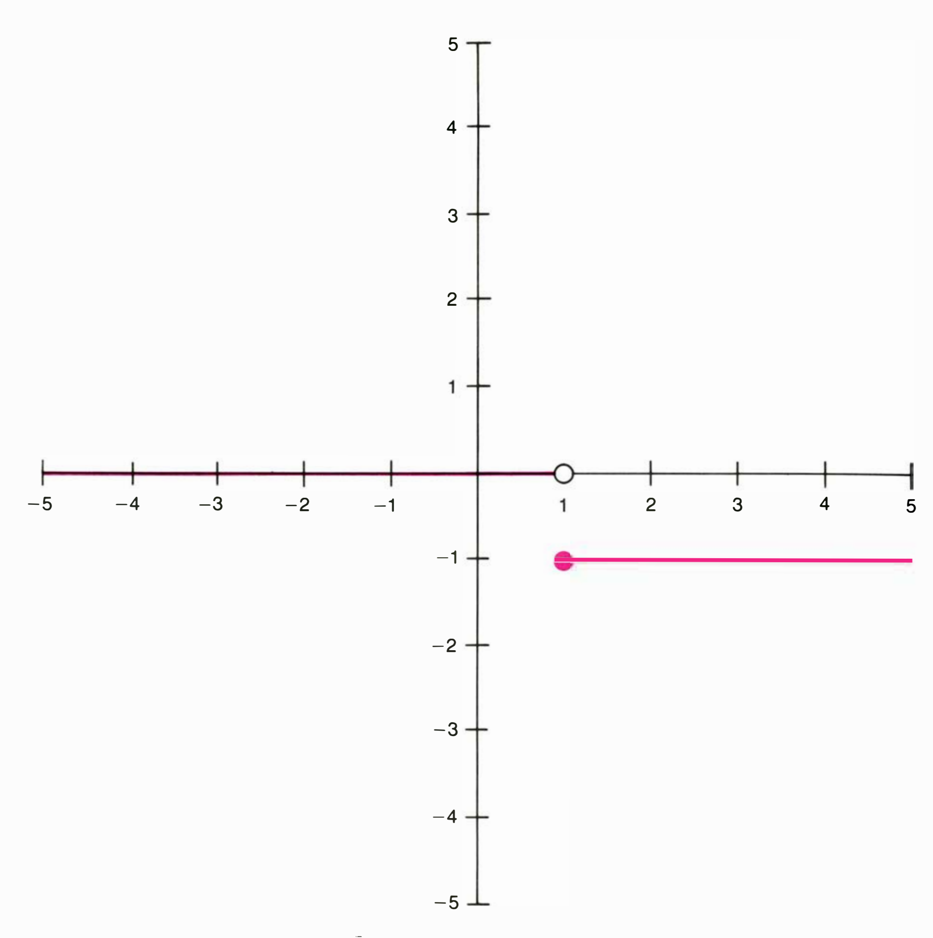


AXIOMA DE ELECCION; se trata de uno de los más ampliamente utilizados y más controvertidos principios de la moderna teoría de conjuntos. Enuncia que, dada una colección cualquiera de conjuntos no vacíos, siempre es posible elegir un representante en cada conjunto de la colección. O sea, dada una colección, como por ejemplo, una de las mostradas aquí, sería posible especificar una regla de selección de representantes tal que dos personas que fueran inspeccionando la colección efectuasen la misma elección de representantes. Por ejemplo, para la colección infinita de pares de zapatos que vemos arriba, la regla "Elíjase el zapato izquierdo de cada par" sería suficiente. Empero, parece que no existirá ninguna regla semejante para la colección infinita de pares de calcetines que se muestra en la parte inferior. Este ejemplo, debido a Bertrand Russell, hace uso de colecciones infinitas de conjuntos; pero incluso para colecciones finitas pueden presentarse problemas parecidos. En la versión de la teoría de conjuntos propuesta por Bishop, estos problemas parecen quedar eliminados, porque, para definir conjuntos constructivistamente, no sólo es necesario dar una regla que permita decidir si un elemento pertenece o no a un conjunto, como pide la teoría de conjuntos tradicional, sino también un procedimiento (finito) para construir un elemento del conjunto. Para demostrar que un conjunto no es vacío, por tanto, es necesario probar que el método de construcción genera efectivamente un elemento del conjunto. En consecuencia, siempre se puede suponer que cada conjunto de la colección contiene un elemento distinguido, a saber, el que fue necesario construir para mostrar que el conjunto no era vacío. El axioma de elección siempre se satisface.

construidos mediante la definición de Bishop obedecen todas las reglas de la aritmética ordinaria, a saber, para cualesquiera dos números x e y que puedan ser construidos mediante esta definición, será posible construir su suma $x+y$ y su diferencia $x-y$, etcétera. Para ver que $x+y$ puede ser aproximado tanto como se desee por números racionales, por ejemplo, en menos de $1/10^n$, observemos que por ser posible construir x e y, han de existir números racionales r y s que difieran de x e y, respectivamente, en menos de $1/10^{n+1}$. Por consiguiente, su suma $r+s$ difiere de $x+y$ en menos de $2/10^{n+1}$, que es menor que $1/10^n$. En consecuencia, $r+s$ es aproximación adecuada.

Una vez establecido un procedimiento adecuado de construcción de los números reales, para completar la definición del conjunto de estos números falta solamente dar un procedimiento que permita decidir si dos números reales son distintos. Lo que se consigue muy fácilmente, pues basta declarar distintos a dos números reales, x e y, siempre que sea posible construir un número racional r situado entre ambos. Dicho de otra forma, o bien x habrá de ser mayor que r y r mayor que y, o bien x habrá de ser menor que r y r menor que y.

Una curiosa consecuencia de la noción constructiva de conjunto dada por Bishop es que ya no es posible descomponer el conjunto de los números reales en dos conjuntos no vacíos y sin elementos comunes; descomposiciones frecuentemente utilizadas en la matemática contemporánea. Por ejemplo, en la teoría de conjuntos no constructivista, el conjunto G formado por todos los números reales menores de 1, reunido con el conjunto D formado por todos los números reales mayores o iguales que 1, puede identificarse con el conjunto completo, formado por todos los números reales. Sin embargo, desde el punto de vista constructivista no sucede así, porque no es posible asignar ciertos números reales, como el $a+b$ antes descrito, ni a un conjunto ni al otro. Y tampoco podemos decir que $a+b$ pertenece o bien a G o bien a D, y que, por tanto, la reunión de G con D contiene a todos los números reales, pues ello exige aplicar el principio del tercero excluido al conjunto infinito de todos los números reales. (Siempre existe la posibilidad de que números como $a+b$, definidos con auxilio de un problema aritmético no resuelto, asignen a uno de los conjuntos G o D. Y parece ser hipótesis segura que siempre habrá en la aritmética problemas no resueltos, con lo que podrán definirse números parecidos.)



LAS FUNCIONES ESCALONADAS (*líneas de color*) **forman una clase muy utilizada de funciones reales de variable real que no pueden ser definidas constructivistamente. En matemática no constructivista, una función real de variable real es una regla que a cada número real r le asigna otro número real, $f(r)$. Por ejemplo, la regla que define la función escalonada de la ilustración es: Si r es menor que 1, entonces $f(r)$ es igual a 0; si r es mayor o igual que 1, entonces $f(r)$ es igual a 1. En ciertos casos, sin embargo, es posible definir un número real y a pesar de ello ser imposible determinar en número finito de pasos si tal número es menor que 1 o mayor o igual que 1. En consecuencia, no hay forma de determinar si a tal número debe asignársele el valor 0 o el valor 1. En la matemática constructivista de Bishop se requiere que la regla que defina a cada función real sea computable de manera efectiva. O sea, para todo r ha de poderse construir en número finito de pasos el valor $f(r)$. Por tanto, la función escalonada no puede ser definida constructivistamente sobre todo el conjunto de los números reales. Bishop predijo que tan pronto se viesen las ventajas de su programa, la matemática moderna dejaría de existir.**

La anterior restricción sobre descomposiciones del conjunto de los números reales se refleja en los tipos de funciones de variable real que pueden ser definidas constructivistamente. En la teoría de conjuntos no constructivista, una función f de un conjunto S en un conjunto T es una regla que asigna a cada elemento s de S un único elemento t de T, es decir, $f(s)$ es igual a t. Por ejemplo, la regla que a cada persona le asigna su madre puede considerarse como una función del conjunto de todas las personas en el conjunto de todas las mujeres. Como este ejemplo indica, a cada miembro de S debe asignársele uno y sólo uno de los miembros de T, pero no es necesario que todos los miembros de T estén asignados a algún miembro de S. (O sea, toda persona ha tenido madre, pero no todas la mujeres son madres).

En la teoría de conjuntos constructivista, una función de S a T sigue siendo una regla que a cada elemento de S le asigna un elemento de T, pero se requiere además que tal regla sea calculable de manera efectiva, es decir, que para cada elemento concreto s de S tiene que poderse construir el elemento $f(s)$, o sea, t, que le está asignado. Examinemos ahora la función que va del conjunto de números reales en sí mismo determinada por la siguiente regla: $f(r)$ es igual a 0 siempre que r sea menor que 1, y $f(r)$ es igual a 1 siempre que r sea mayor o igual que 1. Las funciones de este tipo, llamadas funciones escalonadas por el aspecto que ofrecen sus gráficas, se utilizan ampliamente en la matemática contemporánea [*véase la ilustración de esta misma página*]. Sin embargo, constructivistamente no están bien definidas, porque no es posible decidir si números como el $a+b$ se encuentran en el con-

junto de números menores que 1 (G) o en el conjunto de números mayores o iguales que 1 (D), y por consiguiente no es posible decidir si a tales números debe asignárseles 0 o 1. Por otra parte, tal función, considerada definida no sobre todo el conjunto de los números reales en sí mismo, sino solamente sobre el conjunto formado por G y D (con valores en el conjunto de todos los números reales), sí es perfectamente aceptable.

La función escalonada no constructivista definida sobre el conjunto de todos los números reales es uno de los más sencillos casos de función discontinua. Hablando desde una óptica intuitivista, una función del conjunto de números reales en sí es continua cuando el valor $f(r)$ varía continuamente, sin saltos. (Heurísticamente, la gráfica de una función continua puede dibujarse de un trazo, sin levantar el lápiz del papel.) Conforme r atraviesa el valor 1, el valor $f(r)$ de la función escalonada salta de 0 a 1, y por ello la función es discontinua. Dado que los números reales están definidos con ayuda de los números racionales cercanos a ellos, el valor que una función definida constructivistamente asigne a un número real dado dependerá de los valores que asigne a los racionales cercanos. Por esta razón, parece improbable que sea posible definir constructivistamente funciones sobre todo el conjunto de los números reales que no sean continuas en todos los puntos. (Aunque esta observación, que era válida para la construcción de los números reales dada por Brouwer, no ha podido ser demostrada para la construcción de Bishop, tampoco ha podido encontrarse nunca una función discontinua definida constructivistamente.)

Las funciones discontinuas se presentan constantemente en matemática, particularmente en la aplicada; por ello, la perspectiva de que la teoría constructivista permita solamente definir funciones continuas no es del todo agradable. Por otra parte, ni siquiera con la teoría no constructivista podemos aplicar una función discontinua a todos los números reales. No se puede determinar el valor de, digamos a modo de ejemplo, la función escalonada correspondiente a un número real dado en tanto no se haya determinado en cuál de los conjuntos G o D se encuentra el número. Por consiguiente, desde el punto de vista práctico, esta peculiaridad de las funciones de la teoría constructivista es menos perturbadora de lo que podría temerse a primera vista.

Debido a que las exigencias de la matemática constructivista son más estrictas que las de la no constructivista, todo teorema constructivistamente verdadero también es verdadero en la teoría no constructivista. (Puede suceder, empero, que dos teoremas distintos en la teoría constructivista tengan una misma interpretación no constructivista.) Por consiguiente, podría considerarse que la matemática constructivista es una rama de la matemática moderna. Por otra parte, todo teorema no constructivista puede convertirse en teorema de la teoría constructivista añadiendo a su enunciado "admitido el principio de tercero excluido". O bien, demostraciones no constructivistas pueden ser reinterpretadas desde el punto de vista constructivista. Para fijar ideas, examinemos en esta perspectiva la demostración no constructivista de que existen números irracionales a y b tales que el número a^b es racional.

En la teoría no constructivista, esta proposición se prueba como sigue. El número $\sqrt{2}$ es irracional, y por la ley de tercero excluido, el número $\sqrt{2}^{\sqrt{2}}$ es, o bien racional, o bien irracional. Si $\sqrt{2}^{\sqrt{2}}$ fuese racional la proposición quedaría demostrada sin más que tomar a igual a $\sqrt{2}$ y b igual a $\sqrt{2}$. Por otra parte, si $\sqrt{2}^{\sqrt{2}}$ fuese irracional, la proposición quedaría demostrada tomando a igual a $\sqrt{2}^{\sqrt{2}}$ y b igual a $\sqrt{2}$, porque $(\sqrt{2}^{\sqrt{2}})^{\sqrt{2}}$ es igual a $\sqrt{2}^{\sqrt{2}\sqrt{2}}$, que es igual a $(\sqrt{2})^2$, es decir, a 2. (Esta demostración ofrece particular interés, porque a diferencia de muchas demostraciones de pura existencia casi llega a generar un ejemplo. No obstante, para obtener verdaderamente un ejemplo, sería necesario saber si realmente $\sqrt{2}^{\sqrt{2}}$ es racional, o si es irracional.)

¿Qué manifiesta esta demostración desde una perspectiva constructivista? Para empezar, indica que todo intento de demostrar la proposición de que para todo par de números irracionales a y b también a^b es irracional sería perder el tiempo. Puede demostrarse desde un enfoque cnstructivista que tal tarea está condenada al fracaso suponiendo verdadera la proposición. En tal caso, $\sqrt{2}^{\sqrt{2}}$ debe ser irracional, y por el mismo razonamiento, $(\sqrt{2}^{\sqrt{2}})^{\sqrt{2}}$, o sea, 2, debería ser irracional, contradicción evidente. Por tanto, la demostración establece constructivistamente que no es el caso que para todo par de números irracionales a y b también a^b sea irracional.

Bishop predijo que tan pronto se viesen claramente las ventajas de su programa, la matemática moderna, en la forma que se practica ordinariamente, dejaría de existir como disciplina independiente, pasando a convertirse en rama de la matemática constructivista.

La edición de *The Foundations of Constructive Analysis* está en la actualidad agotada, y en los 12 años transcurridos desde su publicación pocos signos se han producido que muestren que tal transformación se esté efectuando. A pesar de que el número de matemáticos que practican la teoría constructivista está creciendo lentamente, en su mayor parte los matemáticos han visto pocas o ninguna razón para cambiar radicalmente sus métodos fundamentales de trabajo. A pesar de los problemas de fundamentos todavía no resueltos, la matemática moderna ha tenido un éxito extraordinario. Por otra parte, no está claro que la acusación de los constructivistas –que la matemática moderna carece de significado real –sea válida. Después de todo, la utilidad de la matemática en el mundo real es innegable. A quienes trabajan en matemática aplicada, esta disputa, que Einstein calificó de "batalla de ranas y ratones entre matemáticos", les parece totalmente intrascendente. Además, los problemas importantes, los que determinan el desarrollo posterior de la matemática, son por lo general sumamente intuitivos, con lo que la disciplina nunca se aparta excesivamente de la "realidad". La creencia, ampliamente difundida, de que la matemática es, aún en sus versiones más abstractas, mucho más que un juego carente de significado, parece estar bien fundada.

Incluso si nunca llegase a producirse la revolución constructivista, el futuro de la matemática constructivista, en tanto que vigorosa disciplina matemática, parece asegurado. La actual avalancha de producción de resultados constructivistas ha mostrado que, en ciertos casos, afrontar los problemas desde este ángulo reporta ventajas considerables. No solamente pueden conseguirse nuevos, y en ocasiones más profundos, resultados, sino que, por lo general, los teoremas obtenidos son más útiles. Por ejemplo, cuando ha podido demostrarse constructivistamente la existencia de un cierto número, es frecuente que tal número pueda calcularse no sólo en principio, sino efectivamente en la práctica. Bishop ha hecho notar que la moderna teoría de probabilidad dispone de pocos métodos generales para computar probabilidades de manera efectiva, y por ello se está llevando a cabo el desarrollo constructivista de esta disciplina, principalmente por un antiguo discípulo de Bishop, Y. K. Chan. La matemática constructivista se asignará, con provecho, a muchas ramas de la matemática aplicada, donde, a fin de cuentas, es mucho más importante conseguir soluciones de los problemas que demostrar simplemente que éstas existen.

La mayoría de los actuales matemáticos son ya producto de varias generaciones de educación matemática formalista, lo que los coloca en una "ratonera" mental, desde la cual es difícil observar con objetividad la matemática. Desde el interior de esta "trampa", la única manera de poner en tela de juicio la certidumbre de la matemática sería descubrir incoherencias inevitables e inherentes al sistema formal axiomático de la matemática. Y es opinión general que, si bien tales incoherencias pudieran existir, podrían corregirse retocando ligeramente los axiomas. De esta forma, la cuestión de si la matemática tiene significado real ni siquiera puede plantearse. Y respecto a los matemáticos que pudiesen dudar de la validez de su disciplina, la opinión más difundida es que deberían seguir el consejo de Jean d'Alembert, filósofo francés del siglo XVIII, a quienes dudaban del cálculo diferencial: "Persevere, y la fe volverá".

El enfoque formalista, que no autoriza más crítica que la formulada en sus propios términos, ha hecho caer en excesos a los dos bandos de la polémica constructivismo-formalismo. En concreto, la gratuita convicción de que el constructivismo es un cáncer que terminaría destruyendo el edificio matemático ha impedido analizar sistemáticamente la noción de existencia tangible. Puede ser imposible convertir en demostraciones constructivistas todas las demostraciones de existencia pura, pero frecuentemente sí sería posible partir de puntos donde la existencia tangible haya sido demostrada ya, y construir sobre esta base. Por otra parte, los constructivistas descarrían cuando rechazan categóricamente la noción de existencia pura, que sí implica la noción, constructivistamente válida, de que sería contradictorio para algo que es, no existir.

Es admitido por todos que la capacidad matemática procede de la habilidad de la mente humana para combinar razón e intuición. No obstante, la vitalidad y vigor de la matemática emana sobre todo de la intuición. Si el constructivismo recibiese consideración adecuada, sin duda se establecería un más correcto reparto entre razón e intuición. Además, la creencia de que la verdad matemática existe con independencia del conocimiento humano es acto de fe que muchos matemáticos han realizado inadvertidamente. Los constructivistas están comenzando a demostrar que gran parte del trabajo matemático puede llevarse a cabo prescindiendo de tal hipótesis.

Juegos matemáticos

Ciertos problemas de empaquetado no se resuelven sentándose en la maleta

Martin Gardner

"¡*Es* un cuadrado!... ¡Qué hermoso!¡Cuán her-mo-so! ¡Equilátero! ¡*Y* rectangular!"

—LEWIS CARROLL,
A Tangled Tale, Knot Two

En la sección correspondiente al número de marzo analicé el problema, no resuelto en el caso general, de alojar n círculos idénticos en cuadrados de área mínima, sin que los círculos se superpongan. El empaquetamiento de n cuadrados idénticos en "marcos" cuadrados de superficie mínima presenta problemas parecidos, tampoco resueltos salvo para pequeños valores de n. Este mes, tras recapitular los dispersos resultados que se conocen acerca de esta difícil tarea, me fijaré en algunas otras curiosas cuestiones relativas a empaquetamientos de cuadrados, y terminaré con una breve incursión en el campo de las tres dimensiones.

Lo mismo que ocurría al envasar círculos en cuadrados, la tarea de alojar cuadrados en marcos cuadrados puede enfocarse desde dos ángulos. Por una parte, si se supone que el lado del cuadrado exterior es de longitud 1, el problema consiste en determinar cuán grandes podrán hacerse los cuadrados iguales para poder todavía alojar n de ellos en el cuadrado exterior. Por otra parte, si se supone que los cuadrados iguales sean todos de lado 1 (en cuyo caso los llamaremos cuadrados unitarios), el problema consiste en determinar el mínimo cuadrado capaz de enmarcar los n cuadrados. Este último enfoque será el que adoptemos aquí. Y aunque parezca increíble, solamente han podido encontrarse soluciones minimales cuando n sea entero cuadrado perfecto, o sea igual a 2, 3 o 5.

Llamaremos k al lado del cuadrado que habrá de enmarcar los cuadrados unitarios. La ilustración de la página siguiente muestra los empaquetamientos óptimos conocidos (es decir, los correspondientes a valores mínimos de k) cuando el número de cuadrados unitarios va desde 1 hasta 15; cada empaquetamiento se comenta brevemente. Desde luego, siempre que n sea cuadrado perfecto, k será la raíz cuadrada de n; y en los casos donde n sea 2, 3 o 5 no es difícil demostrar que los valores de k dados en la ilustración son mínimos. Cuando n no sea cuadrado perfecto, el valor de k deberá ser mayor que la raíz cuadrada de n, y cuando n sea pequeño, k será menor que el mínimo entero mayor que la raíz cuadrada de n.

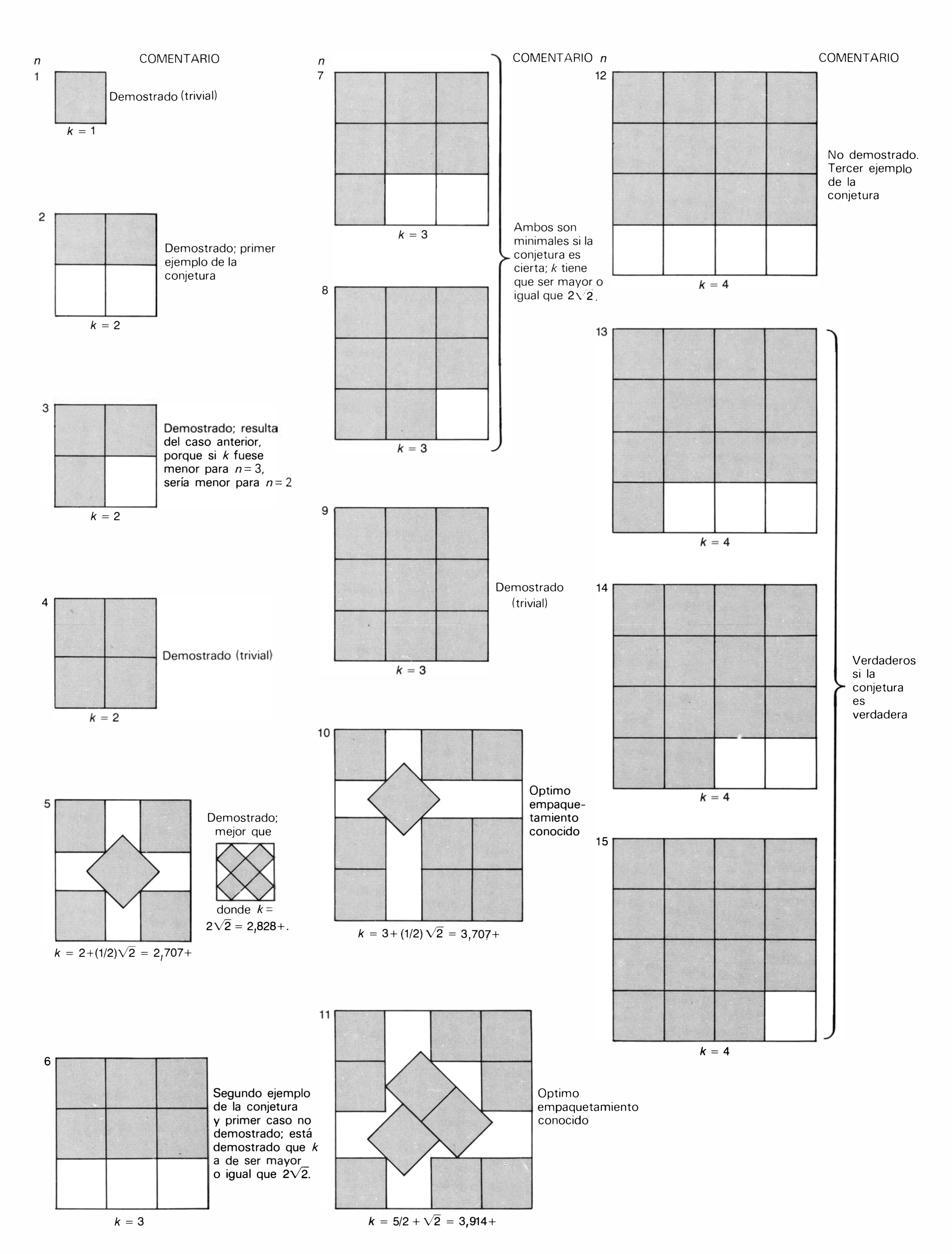
Se ha propuesto como conjetura que si n no es demasiado grande, cuando n sea igual al valor $a^2 - a$ para algún número entero a, el valor de k, lado del mínimo cuadrado capaz de albergar n cuadrados unitarios, será precisamente a. Podemos enunciar la conjetura en lenguaje geométrico: los cuadrados unitarios de una formación de $(a-1)$ por a no podrán reagruparse de forma que encajen en un cuadrado de lado más pequeño que a. De ser cierta tal conjetura, resultaría inmediatamente que un cuadrado del mismo tamaño sigue siendo minimal cuando se hace crecer el número de cuadrados unitarios, hasta el valor a^2 inclusive. Para comprenderlo, supongamos que el cuadrado marco pudiese tener lado menor que a, y dar todavía acomodo a uno o más cuadrados aprovechando el espacio correspondiente a la fila vacía (que podemos imaginar es la a-ésima). Al retirar el cuadrado —o cuadrados— extra nos quedarían $a^2 - a$ cuadrados unitarios contenidos en un marco de lado menor que a, en contradicción con la conjetura.

Se ignora cuál sea el mínimo valor de n para el que la conjetura deja de ser verdadera, pero Ronald L. Graham, de los Laboratorios Bell, ha comprobado que la conjetura cae en falta cuando n vale $40^2 - 40$ (es decir, cuando el cuadrado exterior mide 40 de lado), pero en su opinión el mínimo valor de n para el que la conjetura ya no es cierta debe ser considerablemente menor, quizá, del orden de $10^2 - 10$ (o sea, 90). Tampoco se conoce el mínimo valor de n para el cual se precisa sesgar los cuadrados con ángulos distintos de 45 grados.

Los resultados anteriores, lo mismo que los presentados en la ilustración de la página adyacente, se han tomado de la única fuente que conozco al respecto de este fascinante problema: un artículo titulado "Geometrical Packing and Covering Problems", del matemático holandés F. Göbel. Evidentemente, cuando n es igual a 16 el valor de k es 4. Los empaquetamientos correspondientes a n igual a 17 y 18 que dan los mejores valores de k conocidos hasta el momento se muestran en la parte superior de la ilustración de la página 112. Si la conjetura antes mencionada fuese cierta, cuando n estuviera entre 20 y 25 (ambos inclusive) el valor de k sería 5.

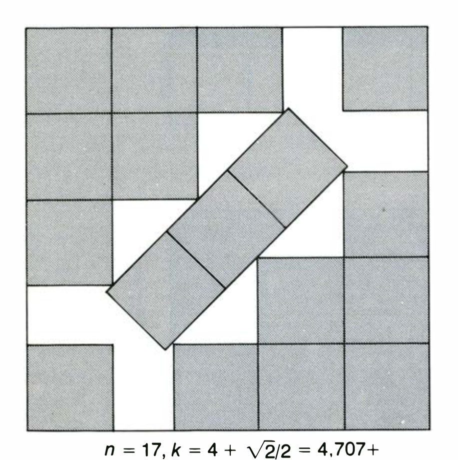
El caso $n = 19$ ofrece particular interés. A principios de este año, tras leer en esta sección el artículo sobre empaquetamientos de círculos unitarios en cuadrados, Charles F. Cottingham comenzó a estudiar empaquetamientos de cuadrados unitarios en cuadrados mayores. Desconocedor de trabajos anteriores sobre el tema, Cottingham duplicó los resultados de Göbel hasta $n = 25$, con sólo una excepción. En el caso $n = 19$ Cottingham obtuvo un empaquetamiento ligeramente mejor que el mostrado en la parte inferior de la ilustración de la página citada. ¿Sabrá el lector mejorar esta colocación antes del mes que viene, en que daré la solución de Cottingham? Cottingham perfeccionó también algunos empaquetamientos de Göbel para ciertos valores de n mayores que 25.

Conforme n va creciendo, la tarea de probar que un marco cuadrado es minimal (caso de no ser n cuadrado perfecto) se hace más y más difícil. Ni siquiera con $n = 5$ resulta trivial la demostración. Un procedimiento para determinar la mínima longitud del lado del cuadrado que encierra cinco cuadrados unitarios consiste en dividir el cuadrado-marco en cuatro regiones cuadradas iguales. Como dentro han de alojarse cinco cuadrados, el principio del casillero asegura que al menos dos de los centros de los cuadrados unitarios han

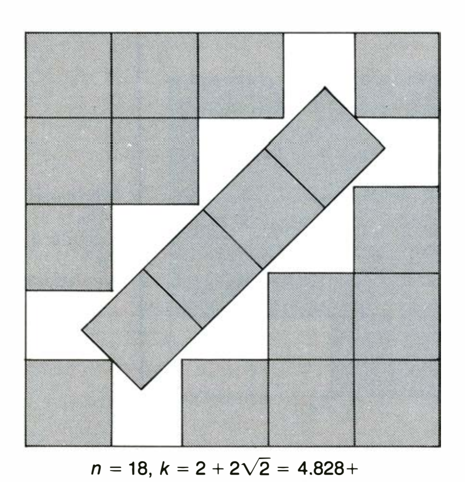


Resultados óptimos conocidos para alojar **n** *cuadrados unitarios en el mínimo cuadrado posible, para* **n** *de 1 hasta 15*

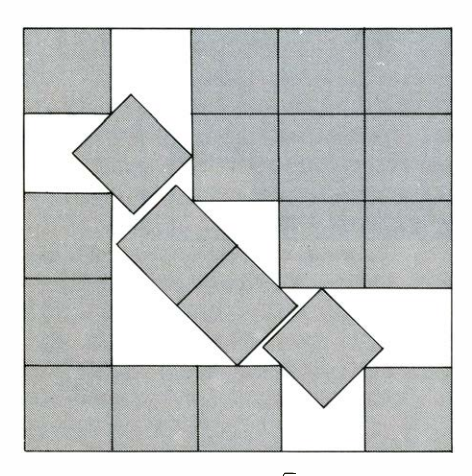
de encontrarse dentro, o en el borde de una (al menos) de las cuatro regiones. El paso final consiste en demostrar que una región lo suficientemente grande como para alojar los centros de dos cuadrados unitarios no puede tener lado menor que $1 + (1/4)\sqrt{2}$. Por consiguiente, el lado del cuadrado exterior minimal ha de ser doble de tal longitud, o sea, $2 + (1/2)\sqrt{2}$.



$n = 17, k = 4 + \sqrt{2}/2 = 4{,}707+$



$n = 18, k = 2 + 2\sqrt{2} = 4.828+$



$n = 19, k = 4 + (2/3)\sqrt{2} = 4{,}942+$

Empaquetamiento para **n** = *17*, **n** = *18 y* **n** = *19*

No hay en perspectiva ningún procedimiento general para determinar marcos cuadrados mínimos, pero en 1975 Paul Erdös y Graham publicaron un artículo donde probaban un notable teorema. Demostraron que, si aumenta el número de cuadrados unitarios hasta un número suficientemente grande, mediante empaquetamientos bien diseñados se puede reducir la proporción de espacio desperdiciado, hasta una superficie que nunca rebase de $k^{7/11}$, o sea $k^{0.636+}$, siendo k el lado del cuadrado exterior. (La superficie desperdiciada es nula cuando el número de cuadrados unitarios es cuadrado perfecto.)

Posteriormente, Hugh Montgomery consiguió rebajar un poco esta cifra –que es cota asintótica–, dejándola en $k^{(3-\sqrt{3})/2}$ o sea, $k^{0.633+}$, y probablemente pueda rebajarse todavía más. Graham ha conjeturado que conforme n tiende a infinito, la cota final puede muy bien ser $k^{0.5}$, o sea, la raíz cuadrada de k. En un artículo publicado en 1978, Klaus F. Roth y Robert C. Vaughan, dos matemáticos ingleses, demostraron que la cota no puede tener límite menor que $k^{0.5}$. Por tanto, en el estado actual del problema puede afirmarse que en empaquetamientos óptimos de gran número de cuadrados unitarios la superficie desperdiciada será mayor que $k^{0.5}$ y menor o igual que $k^{0.633+}$.

Para resaltar la gran cantidad de espacio que puede ahorrarse cuando el número de cuadrados unitarios es grande, Graham examina el empaquetamiento de cuadrados unitarios en un gran cuadrado de lado 100.000,1. Si los cuadrados se alojan, como es tradicional hacerlo, en filas rectas, pueden acomodarse 100.000^2 cuadraditos unitarios. Este empaquetamiento deja una superficie vacía, una orla lineal y extraordinariamente fina, en la que no puede darse cabida a ningún cuadrado de lado unidad. Pero si los cuadraditos se ladean y sesgan según la técnica de Erdös y Montgomery, todavía se puede introducir la friolera de más de 6400 cuadraditos unitarios adicionales.

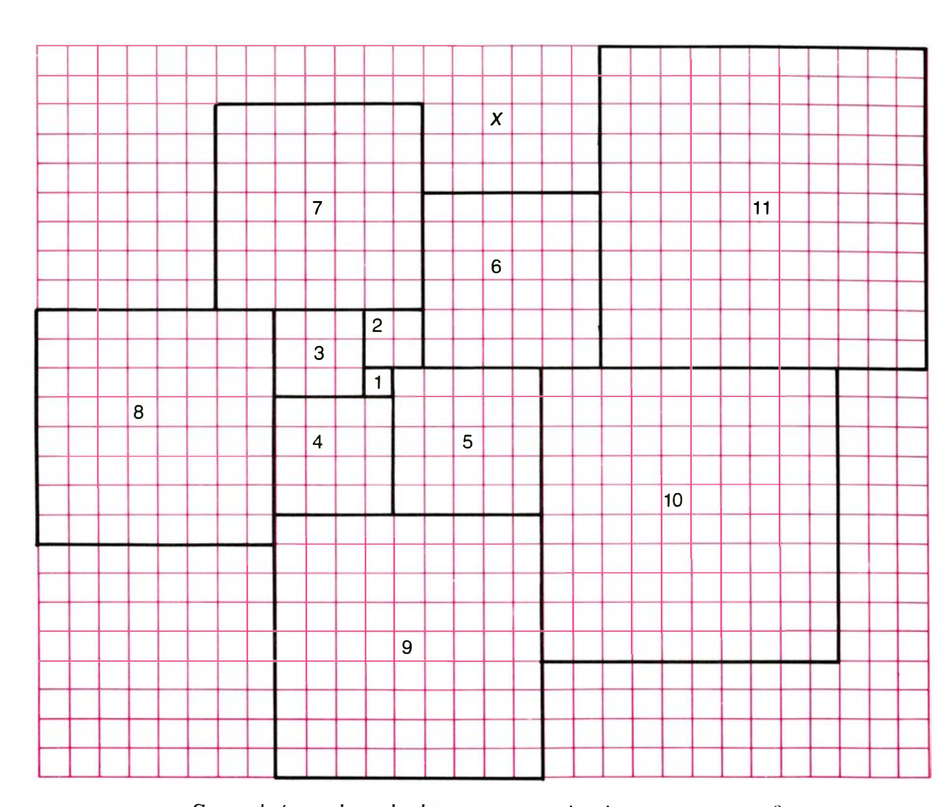
Fijémonos ahora en la sucesión de cuadrados de lados 1,2,3,4,5, ... Las áreas de estos cuadrados forman la sucesión infinita 1, 4, 9, 16, 25, ..., cuyas sumas parciales correspondientes son 1, 5, 14, 30, 55, ... ¿Habrá en esta sucesión de sumas alguna que sea cuadrado perfecto? Pues sí, pero solamente una: 4900. A mediados de la década de 1960, este número, que es cuadrado de 70 y suma de los primeros 24 cuadrados de números naturales, suscitó el siguiente problema: ¿Podrá el cuadrado de 70 por 70 ser rellenado exactamente con los primeros 24 cuadrados de lados enteros consecutivos? La respuesta a esta cuestión resulta ser negativa, pero hay formas de empaquetar casi todos los cuadrados de la colección –omitiendo el cuadrado de 7 por 7– que dejan tan sólo 49 unidades cuadradas desperdiciadas. [Véase mi *Mathematical Carnival*, p. 149.] Se admite que tal empaquetamiento es minimal, pero yo no conozco ninguna demostración de esta conjetura. No sé tampoco si podría conseguirse un rellenado con menos espacio desocupado imaginando que el "cuadrado" de orden 70 forma una superficie cilíndrica o tórica.

¿Se podrá pavimentar un rectángulo con una colección de cuadrados de lados enteros consecutivos (ya empezando por lado igual a 1, ya empezando más adelante)? Este problema, original de Solomon W. Golomb está pendiente todavía de solución. De ser negativa la respuesta, como muchos sospechan, Golomb sugiere el siguiente refinamiento: Al alojar cuadrados de lados consecutivos en un cuadrado más grande, ¿qué disposición, empezando por 1, dejará un tanto por ciento mínimo de espacio desperdiciado? La misma pregunta puede plantearse para marcos rectangulares. Que yo sepa, nada se ha hecho hasta el momento para responder a ninguna de estas cuestiones.

Ningún cuadrado puede ser rellenado exactamente con cuadrados de lados 1, 2, 3, 4, 5, ..., y probablemente tampoco pueda pavimentarse ningún rectángulo, pero, ¿será posible teselar el plano (infinito) con cuadrados consecutivos, a partir del de lado unidad? Este insólito problema ha sido propuesto por Golomb en el *Journal of Recreational Mathematics* (vol. 8, n.º 2, págs. 138-139; 1975-76). Con los cuadrados consecutivos de lados 1 hasta 11 puede construirse una espiral, como la que vemos en la ilustración de la página siguiente, pero tal pavimentación deja inevitablemente un hueco, el espacio señalado por *X*. Todavía no se ha descubierto ningún algoritmo para construir pavimentaciones completas del plano, ni siquiera comenzando por algún entero adecuado, aunque tampoco se ha demostrado que la tarea sea imposible.

La ilustración superior de la página 114 muestra una forma de ir colocando cuadrados de lados consecutivos que constituyan una región simplemente conexa (una región sin agujeros) que recu-

bra más de las tres cuartas partes del plano. El cuadrante sureste está cubierto por baldosas cuyos lados forman la conocida sucesión de Fibonacci: 1,2,3,5,8,13, ... El cuadrante suroeste, por una sucesión de Lucas truncada, que comienza por 7,11,18,29 ... más una delgada banda de solapamiento con los cuadrados situados a su derecha. (La sucesión de Lucas, que comienza 1, 3, ... es la más sencilla de las sucesiones generalizadas de Fibonacci.) El cuadrante noreste está cubierto por cuadrados en sucesión de Fibonacci truncada, que comienzan 6, 9, 15, 24, ... más una delgada tira producida por el solapamiento con los cuadrados situados debajo. Verner E. Hoggatt, Jr., editor de *The Fibonacci Quarterly*, ha demostrado que no hay duplicidad numérica en estas tres sucesiones. Los cuadrados de lados no utilizados en ninguna de estas sucesiones están dispuestos sobre los cuadrados de la sucesión de Lucas. El punto negro de la ilustración es el origen del plano cartesiano. Esta configuración de cuadrados consecutivos recubre más de tres cuartas partes del plano, como es obvio, pero se desconoce si tan siquiera se podrá pavimentar exactamente un cuadrante del plano mediante cuadrados de lados consecutivos.



¿Se podrá teselar el plano con cuadrados consecutivos?

Existen muchas formas de pavimentar el plano completo si se exige, no que los cuadrados tengan lados consecutivos, sino solamente que sean todos distintos. He aquí una de ellas, que se apoya en una disección del cuadrado en tan sólo 21 cuadrados menores, todos ellos distintos. (Puede verse una reseña de tan reciente descubrimiento en la sección "Ciencia y sociedad", de *Investigación y Ciencia*, agosto de 1978.) El cuadrado mínimo de una disección así ocupa forzosamente posición interior, y por tanto, para teselar todo el plano tomaremos un cuadrado descompuesto en 21 subcuadrados, que servirá como cuadrado interior mínimo de una copia a mayor escala de la misma configuración. El cuadrado grande teselado de este modo se usa entonces como cuadrado mínimo de una copia todavía mayor, y así sucesivamente. Prosiguiendo de esta forma, se puede recubrir completamente el plano infinito mediante cuadrados de tamaños todos distintos.

Desde hace más de una década se sabe que cualquier conjunto de cuadrados cuya área total sume 1 (no es preciso que sus lados sean números racionales) puede ser alojado sin solapamientos en un cuadrado de área 2 [véase *Mathematical Carnival*, capítulo 11]. Imaginemos un rectángulo de anchura unidad. ¿Cuál es la longitud mínima necesaria para que el rectángulo pueda alojar cualquier colección de cuadrados cuya suma total de áreas sea igual a 1? Desde hace mucho se suponía que la solución era $\sqrt{3}$, conjetura que verificaron Daniel J. Kleitman y Michael M. Krieger en su artículo "Packing Squares in Rectangles I", publicado en *Annals of the New York Academy of Sciences* (vol. 175, artículo 1, págs. 253-262; 31 de julio de 1970). Posteriormente, en "An Optimal Bound for Two Dimensional Bin Packing" (*16th Annual Symposium on Foundations of Computer Science*, IEEE Computer Society, 1975) estos mismos dos autores demostraron que el rectángulo de área mínima en que podría alojarse toda colección de cuadrados de área total igual a 1 tiene dimensiones $2/\sqrt{3}$ y $\sqrt{2}$. En uno y otro caso el despilfarro de superficie es enorme, exactamente del orden del 60 por ciento del área conjunta de la colección de cuadrados.

Un problema clásico de empaquetamiento de cuadrados, conocido como "edredón de la señora Perkins", recibe su nombre de un rompecabezas mencionado en uno de los libros de Henry Ernest Dudeney. (En el capítulo 11 de *Mathematical Carnival* se expone también el trabajo de John Horton Conway sobre este problema.) La tarea consiste en dividir un cuadrado de lado entero n en el mínimo número posible de subcuadrados no rampantes y de lados enteros y sin que se desperdicie ningún espacio. En este problema no es preciso que los subcuadrados sean todos distintos.

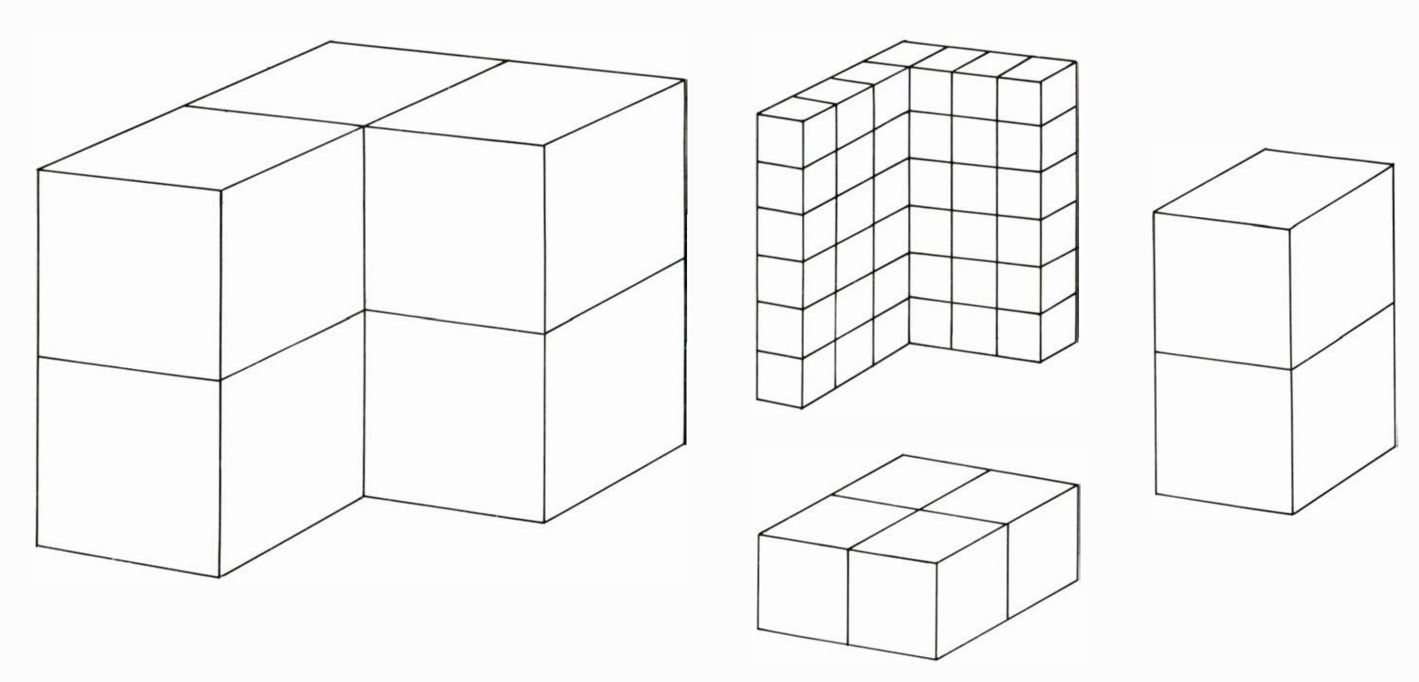
La siguiente cuestión está íntimamente relacionada con el edredón de la señora Perkins. ¿Cuál es el máximo número de subcuadrados (se permiten duplicaciones) en el que no podrá descomponerse ningún cuadrado? Dicho de otra forma, ¿cuál es el mínimo número de subcuadrados (número que llamaremos n) tal que un cuadrado puede dividirse en n (y en cualquier número mayor) de subcuadrados? No es difícil demostrar que un cuadrado no puede descomponerse ni en dos, ni en tres, ni en cinco cuadrados, pero sí en seis, o en cualquier número mayor que seis de subcuadrados. (Para dividir un cuadrado en seis subcuadrados, se descompone un cuadrado de 3 por 3 en nueve subcuadrados unitarios, y después se apartan un cuadrado de lado 2, y cinco cuadrados iguales, de lado 1.) Por consiguiente, la respuesta a la primera versión de la pregunta es 5, y con el enunciado de la segunda versión, 6.

La generalización natural de este problema a cubos resultó considerablemente más difícil. Se pudo establecer que un cubo puede subdividirse en n subcubos cuando n es igual a 1, 8, 15, 20, 22, 27, 29, 34, 36, 38, 39, 41, 43, 45, 46, 48, 49, 50, 51, 52, 53, 55 y todos los números superiores. Se demostró

que la tarea propuesta era imposible para todos los restantes valores de n, a excepción de uno: $n = 54$. Durante muchos años, la cuestión de si un cubo podría despiezarse en 54 subcubos permaneció sin resolver, con general perplejidad.

En tanto el caso $n = 54$ estuvo abierto, fue imposible responder la siguiente pregunta: ¿cuál es el máximo número de subcubos, no necesariamente distintos, en que es imposible descomponer un cubo? Tal cuestión, conocida como problema de Hadwiger por haber sido planteada por el matemático suizo Hugo Hadwiger, de Berna, se pensaba que tenía solución 54, pero hace algunos años se encontró una disección en 54 subcubos, por dos residentes en Berna, Dorys Rychener, profesora de flauta, y A. Zbinden, de la International Business

Más de tres cuartas partes del plano recubiertas con cuadrados consecutivos



Disección del cubo de 8 por 8 por 8 en 54 subcubos, original de Doris Rychener

Machines Corporation, independientemente uno de otro. La solución fue publicada por Richard K. Guy en su sección de problemas a investigar en *The American Mathematical Monthly* (vol. 84, n.º 10, p. 810; diciembre de 1977). La ilustración inferior de la página anterior, basada en una ilustración suministrada por Guy, muestra 42 cubos de lado 1, cuatro de lado 2, dos de lado 3, y seis de lado 4. Las cuatro piezas que se ven son ensambladas del modo evidente y producen un cubo de lado 8.

Ahora podemos responder a la cuestión de Hadwiger. El máximo número de cubos en los que es imposible descomponer un cubo es 47; o asimismo, 48 es el mínimo número tal que un cubo puede ser despiezado en dicho número de subcubos, y también en todo número mayor de ellos. Incidentalmente, no es posible descomponer un cubo en ningún número finito de subcubos todos distintos. La demostración, una de las más bellas de la geometría combinatoria, puede verse en mi apéndice al artículo de William T. Tuttle "Squaring the Square", reproducido en mi *2nd Scientific American Book of Mathematical Puzzles & Diversions*.

Muchos curiosos problemas de empaquetamiento de cuadrados se refieren únicamente a un solo cuadrado que es preciso alojar en una región de forma previamente dada. Por ejemplo, ¿cuál es el máximo cuadrado que puede inscribirse en un pentágono regular de lado 1? La solución de este problema, poco conocido y decepcionantemente simple, se dará el mes que viene.

El mes pasado se pidió a los lectores que hallasen una solución de la ecuación $x^y = y^x$, que tuviese x mayor que y, y donde ambos números fuesen enteros y distintos de 4 y 2, que son la solución más conocida. La segunda solución es −2, −4.

Otro ejercicio consistía en demostrar la imposibilidad de que los desarrollos decimales en pi y *e* estén completamente contenidos cada uno en el otro. En efecto, si ello fuese posible, entonces, tras un número (x) finito de cifras del desarrollo decimal de pi comenzaría el desarrollo de *e*, y tras un número finito (y) de cifras decimales de *e* volvería a comenzar pi. En tal caso, tanto *e* como pi tendrían desarrollos decimales periódicos, ambos de período $x + y$. Pero tales desarrollos son característicos de los números racionales, que son los únicos números reales con desarrollos decimales periódicos. Como es sabido, tanto pi como *e* son irracionales. Luego la hipótesis inicial tiene que ser falsa.

Taller y laboratorio

El "rattleback", extraña piedra que primero gira en un sentido y después en el contrario

Jearl Walker

Si se hace girar un cierto tipo de piedra en un sentido "erróneo", se parará muy pronto, se agitará hacia arriba y hacia abajo durante unos segundos y tomará luego un giro en sentido opuesto. Pero si se hubiera puesto a girar en el rumbo "correcto", daría vueltas de forma estable. La piedra está aparentemente inclinada hacia una dirección de giro. Incluso llegará a desarrollar un sentido de giro en esa dirección con sólo golpear levemente un extremo de la piedra hacia abajo. El balanceo de la misma causado por el golpecito se convertirá rápidamente en giro.

Estas piedras tan curiosas fueron llamadas "celtas" originalmente, porque su comportamiento fue descubierto por arqueólogos que estudiaban las hachas y azuelas del mismo nombre. Yo me imagino que el cambio del sentido de giro fue observado primero por un arqueólogo que movía distraídamente una de las "celtas" sobre una mesa. Mi encuentro con estas piedras se produjo hace unos años cuando leí el delicioso libro de Harold Crabtree titulado *Tratamiento elemental del giro de las peonzas y giróscopos*. Entre los diferentes tipos de peonzas, aros y otros juguetes que describía estaban las misteriosas "celtas", pero sin indicar por qué mostraban inclinación hacia una dirección de giro. Mi interés por estas piedras se ha visto reavivado recientemente por dos personas: Nicholas A. Wheeler, del Reed College, que está pensando en la posibilidad de hacer una investigación analítica de las mismas (si llegara a efectuarla, sería el primer estudio de estas piedras desde el siglo XIX, que yo sepa) y A.D. Moore, de la Universidad de Michigan, que me ha enviado algunos de los cientos de "celtas" que ha hecho con piedra, cucharas y otros materiales. Moore ha dado a estas piedras el nombre de "rattlebacks", término que emplearé en el presente artículo.

La mayoría de los "rattlebacks" tienen una base suave elipsoidal, pero la parte superior puede ser plana, hueca o elipsoidal. La inclinación del giro resulta, verosimilmente de la forma de la base y la distribución de la masa con respecto al eje de giro. El rattleback típico de la colección de Moore posee la parte de arriba lisa y rectangular. La característica principal del diseño es que el eje mayor de la elipse forma un ángulo de entre 5 y 10 grados con el eje mayor de la parte rectangular. Esta oblicuidad (y quizás alguna forma sutil de la elipse) introduce la predisposición de la piedra a girar, aunque de una forma que no resulta fácil observarla con detalle.

La mayoría de los rattlebacks que me ha enviado Moore giran, vistos desde arriba, en el sentido de las agujas del reloj. Cuando hago girar un tal rattleback en ese sentido, va rodando suavemente hasta que el rozamiento acaba por detenerlo. Cuando lo hago girar en sentido contrario al de las agujas del reloj, primero da unas cuantas vueltas, pero después se va parando a medida que empieza a balancearse de arriba abajo por sus extremos. Pronto comienza a girar en el sentido horario hasta que lo detiene el rozamiento. Si, cuando está parada, tocamos con un dedo la punta de la piedra, los extremos oscilarán de arriba abajo durante unos segundos y a continuación empezará a girar en el sentido de las agujas del reloj. La piedra se parará, a la larga, debido a la fricción con la mesa.

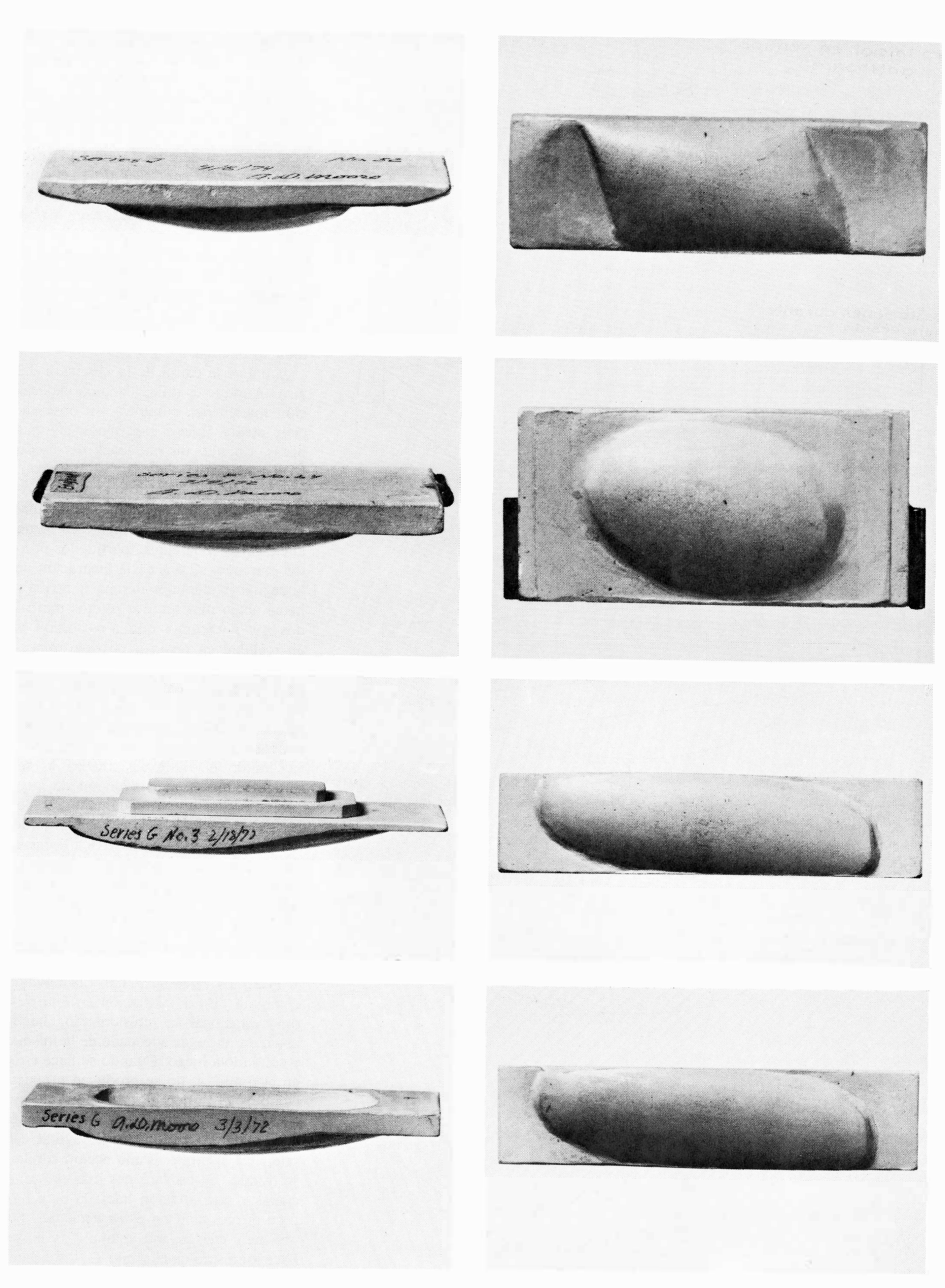
Algunos rattlebacks desarrollan una segunda inversión: pasan de girar en el sentido horario a hacerlo en sentido contrario, poco antes de detenerse justo cuando el rozamiento está a punto de agotar las últimas energías de la piedra. La segunda inversión no suele ser tan fuerte como la primera, ni desarrolla el mismo tipo de oscilación vertical. Antes bien, la piedra se balancea de un lado a otro, no longitudinalmente. Algunos de los rattlebacks desarrollan fuertes inversiones con cualquier sentido de giro.

Moore fabrica la mayoría de sus rattlebacks con cemento dental, al que pule y barniza. Cuando saca uno bueno (trabajo que requiere pericia y suerte) construye un molde del mismo para duplicarlo. Algunos de sus rattlebacks están hechos de la parte cóncava de una cuchara, cuyo mango ha cortado. Después la pega a una pieza plana y rectangular de metal o plexiglás con el eje largo de la elipse de la cuchara en ángulo con el eje largo del rectángulo. A veces añade algunas monedas encima de la pieza rectangular para aumentar el peso y cambiar el momento de inercia del rattleback.

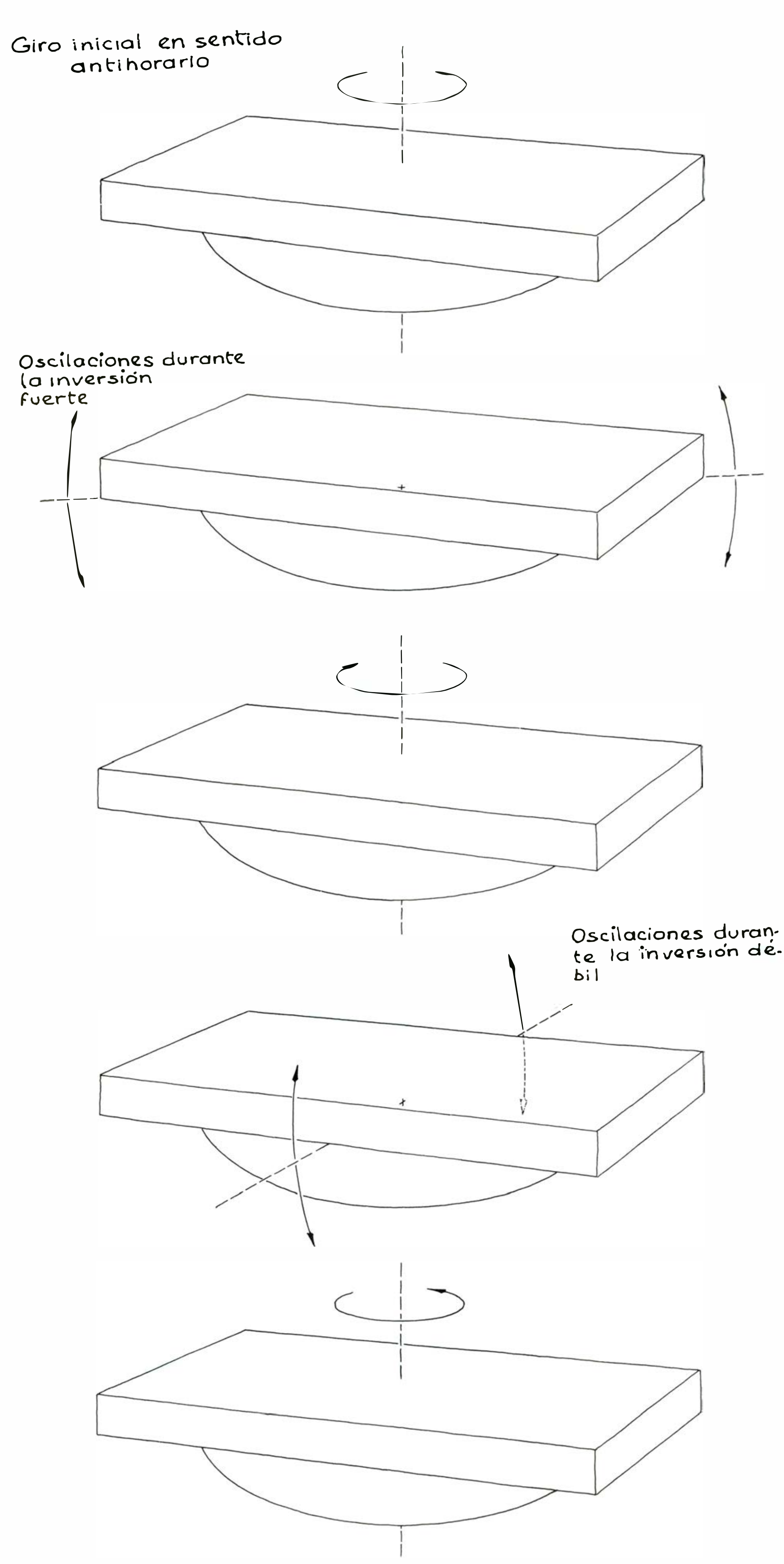
Existe un tipo de rattleback que recuerda la forma del huevo duro cortado por la mitad. Sobre su cara plana, Moore le pegó una barilla de metal en ángulo con el eje largo. Hay otro tipo similar al diseño básico; la única diferencia estriba en que el interior está vaciado para que parezca un barco. Otro modelo se caracteriza por una serie de placas planas y escalonadas en el centro de la parte superior del modelo básico.

El rattleback favorito de Moore debe ser el semihuevo, ya que una vez participó en un concurso con C. L. Stong, autor durante un año de esta sección, para ver cuál de los dos conseguía mayor cantidad de inversiones de giro con un rattleback. Ganó Moore con el modelo de semihuevo, que realizó más de 15 cambios antes de pararse.

Cuando examiné los rattlebacks encontré que el modelo básico mostraba la inversión más fuerte (de girar en sentido antihorario a girar en sentido horario), a la que seguía una segunda inversión más ligera cuando la piedra ya casi había agotado toda su energía. El rattleback con las monedas y el de las placas escalonadas se portaban casi tan bien como el tipo básico. El semihuevo no giraba bien, tendiendo a oscilar tanto que los extremos de la barilla arañaban la mesa y con el roce le restaban fuerza. Moore no se explica bien por qué este modelo, en principio ganador, se porta



Vistas laterales **(izquierda)** *e inferiores* **(derecha)** *de "rattlebacks" hechos por A.D. Moore*



Un "rattleback" que experimenta una inversión fuerte y otra débil

tan mal después, aunque sospecha que la causa podría ser una pequeñísima grieta en la superficie de la base.

El más fascinante de todos los rattlebacks es el que tiene forma de barco, porque desarrolla fuertes inversiones en cada dirección de giro. Cuando lo hago girar en sentido contrario al de las agujas del reloj, efectúa parte de una revolución, se para, oscila bruscamente arriba y abajo e invierte después su giro durante varias revoluciones. Deja de girar otra vez, se balancea de un lado a otro y empieza a dar vueltas en sentido contrario al de las agujas del reloj. Esta actividad continúa hasta que la piedra agota toda su energía.

¿Cuál es la causa de la inversión del giro? Aunque es difícil dar una explicación matemática completa, su observación atenta puede sugerirnos algunas pistas. Cuando un rattleback de diseño básico empieza a invertir su rotación, que tenía sentido contrario al de las agujas del reloj, su oscilación longitudinal parece estar aproximadamente en torno al eje corto de la elipse, porque los puntos marcados *C* y *D* en la ilustración de la página siguiente oscilan hacia arriba y hacia abajo más que los puntos marcados *A* y *F* (creemos que la oscilación es en realidad en torno a otro eje importante; nos referimos a uno de los ejes principales de la piedra, que es casi paralelo al eje corto de la elipse). De alguna forma, la rotación en sentido antihorario provoca una inestabilidad de este tipo.

Cuando el rattleback realiza la segunda y más suave inversión de giro, parece oscilar aproximadamente alrededor del eje mayor de la elipse, enviando los puntos *G* y *H* hacia arriba y hacia abajo con mayor amplitud que otros puntos (de nuevo, teóricamente, el eje real es un eje principal casi coincidente con el eje mayor). Parece ser que el giro hacia la derecha provoca este tipo de inestabilidad.

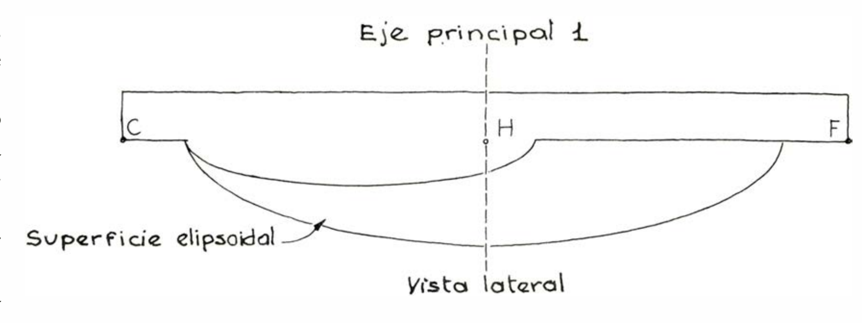
Podemos desencadenar fácilmente cualquier tipo de oscilación en una piedra estacionaria, presionando hacia abajo en un lugar adecuado de la misma y soltándola luego. Cuando se hace esto en el punto *A*, no se producen grandes cambios, el rattleback oscila durante un corto tiempo y girará quizá ligeramente en sentido contrario al de las agujas del reloj. En el punto *B* una acción similar se traduce en oscilaciones más pronunciadas y una rotación hacia la derecha, pero el resultado no es espectacular. Se obtienen efectos más llamativos si se hace lo mismo en el punto *C*; las oscilaciones longitudinales son más vigorosas y la piedra comienza a girar rápidamente en el sentido de las agujas del re-

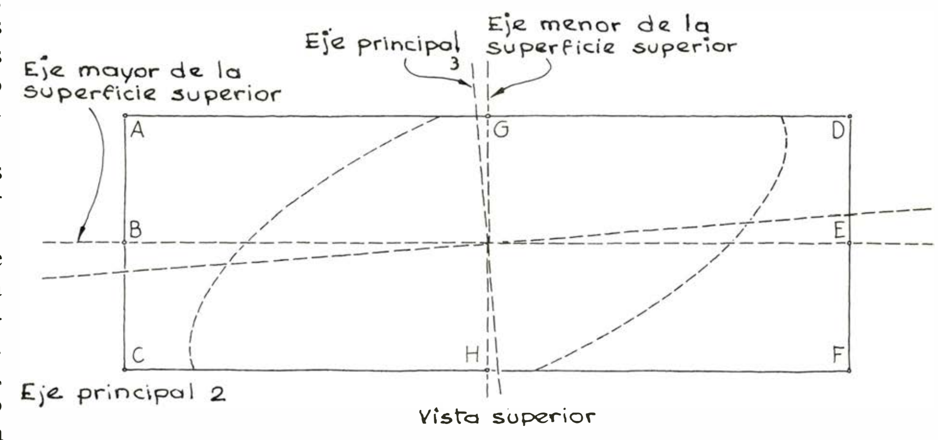
loj. Las oscilaciones son idénticas a las que se pueden ver cuando se produce una inversión en el sentido de giro.

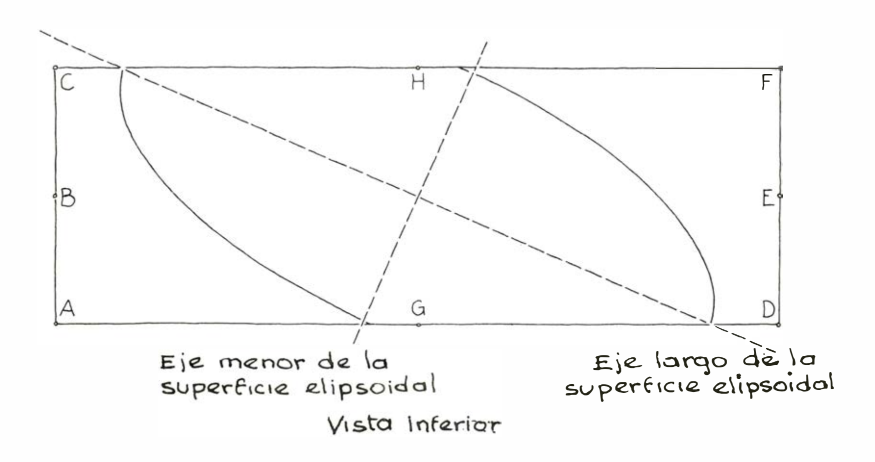
Al presionar, y luego soltar, el punto *G*, la piedra se balancea de un lado a otro y empieza a girar en sentido contrario al de las agujas del reloj. Así pues, cuando se provocan las oscilaciones verticales en *G*, se está duplicando el segundo tipo de inversión de giro (desde la rotación inicial en el sentido de las agujas del reloj). Se obtienen resultados similares cuando se presiona, y se suelta, la piedra en otros puntos entre *G* y *A*. Las oscilaciones transversales resultan más lentas que las longitudinales, pero ambos tipos son tan rápidos que no podemos seguir el movimiento fácilmente. Sin embargo, en cuanto se inician las oscilaciones, la piedra comienza a girar y la perspectiva del observador se altera. Para seguir mejor las oscilaciones traté de manipular un rattleback que había colocado sobre una tira de papel adhesivo con el lado adherente hacia arriba. La cinta evitaba que la piedra girara, pero quizá interfirió en el movimiento natural del objeto. Aun cuando la piedra parecía oscilar principalmente en torno a un eje único, se comportaba también como si rodase fuera del plano primario de oscilación. No estoy seguro de que esta característica sea esencial para la inversión de giro.

La primera investigación científica de piedras que efectuaban inversión de giro se debe a G. T. Walker (que, por cierto, no es ningún antepasado mío), en 1896. Algunos de sus resultados pueden leerse todavía en reimpresiones de libros antiguos relacionados con elementos mecánicos rotatorios. La descripción de las piedras suele darse en el capítulo dedicado a las peonzas asimétricas (peonzas que carecen de simetría respecto al eje vertical). Las discusiones teóricas sobre peonzas simétricas son muy sugestivas, pero la teoría de las peonzas asimétricas constituye un auténtico rompecabezas para la mente.

Un rattleback que se precie debe reunir ciertas características básicas. Hay que atender, por ejemplo, a la no alineación de los ejes principales de la piedra y los ejes del elipsoide de la base. El eje principal de un objeto es aquel en torno al cual se puede hacer girar el mismo libremente sin rotación en torno a ninguno de los otros dos ejes pricipales. Los tres ejes mutuamente perpendiculares se hallan normalmente a lo largo de los ejes de simetría en la mayoría de objetos. Si la elipse estuviera en línea con el extremo rectangular del rattleback básico de Moore, los ejes principales serían fáciles de encontrar, porque coincidirían con los ejes de simetría. Un eje sería vertical y los otros dos serían paralelos a los ejes menor y mayor de la elipse. Todos ellos se cruzarían en el centro de masas de la piedra.







Los ejes de un "rattleback" inclinados para rotar en sentido horario

La base elipsoidal de un buen rattleback no está alineada con la placa rectangular superior. El eje vertical es todavía un eje principal, pero los otros dos han cambiado ahora en torno al vertical en la misma dirección en que se ha sacado la base elipsoidal de la alineación con la placa rectangular. Uno de estos ejes principales, que llamaremos eje 2, está en algún sitio entre el eje mayor de dicha placa rectangular y el eje mayor del elipsoide. El otro, eje 3, se halla entre los dos ejes menores correspondientes y es perpendicular al eje 2. Sus situaciones exactas dependen de la distribución relativa de la masa de la placa rectangular y del elipsoide. La no alineación del elipsoide con los ejes principales es una característica esencial en la inversión del giro.

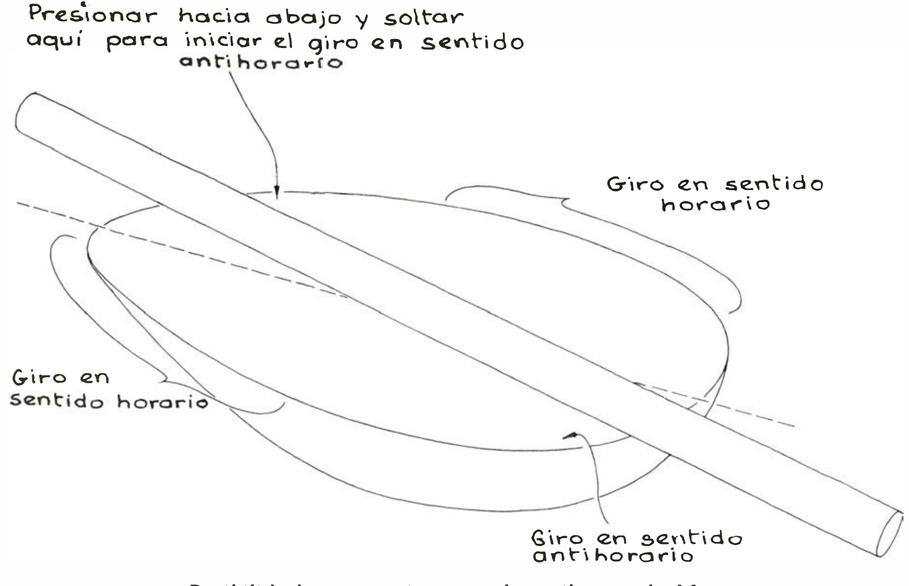
Una segunda característica importante de un rattleback bien hecho es que el radio de curvatura se conserva distinto a lo largo de los dos ejes del elipsoide. (A una superficie que esté poco curvada le corresponderá un gran radio de curvatura). Si la base fuera perfectamente esférica, el rattleback no invertiría su giro.

Importa que la distribución de la masa de la piedra sea diferente para los dos ejes principales. La función que nos da la distribución de la masa con respecto a un eje concreto se llama "momento de inercia". Consideremos el eje principal 3, cuando uno de los rattlebacks de Moore oscila hacia arriba y hacia abajo en torno a ese eje. Como lo hace durante una inversión fuerte de giro, el momento de inercia para la oscilación es relativamente grande porque parte de la masa está a una distancia relativamente grande del eje. Supongamos que la piedra gira en torno al eje principal 2, como lo hace durante el segundo, y más débil, tipo de inversión de giro. El momento de inercia para esa oscilación es relativamente pequeño, porque la mayor parte de la masa está bastante próxima al eje. Esta diferencia en el momento de inercia de los dos ejes horizontales principales es esencial para el éxito del rattleback.

En resumen, el rattleback de Moore tiene las siguientes características: el eje mayor de la elipse (el correspondiente a un radio de curvatura más grande) está desviado en sentido contrario al movimiento de las agujas del reloj (visto desde arriba) del eje principal 2, que tiene un pequeño momento de inercia. Esta disposición causa en el rattleback una fuerte inversión del sentido de giro, que deja de ser antihorario para tener el mismo rumbo que las agujas del reloj y luego una nueva inversión, más débil de la rotación. ¿Qué sucedería si el eje mayor del elipsoide se hubiera desviado en el otro sentido? ¿Se producirían las inversiones de giro fuerte y débil en el sentido opuesto? Aunque el trabajo teórico de G.T. Walker predice el comportamiento, prefiero dejar la pregunta en el aire para que sea el lector quien la responda usando los rattlebacks tallados adecuadamente.

¿Por qué las características de diseño del rattleback de Moore inducen que la piedra cambie su sentido de giro? En primer lugar la hacen inestable frente a las pequeñas perturbaciones del giro alrededor del eje vertical. Si la base fuera esférica, o el elipsoide estuviese alineado con los ejes principales, cualquier pequeña perturbación tendría una escasa repercusión en el giro y no causaría oscilaciones notables. Sin embargo, con una piedra de la forma adecuada, las pequeñas perturbaciones del giro inicial de la piedra o las producidas por el tablero de la mesa generan oscilaciones que crecen rápidamente en amplitud. El hecho de que se desarrolle una inestabilidad u otra depende de la dirección de giro y de la forma de la piedra.

Supongamos que un rattleback con la forma básica está girando en sentido contrario al de las agujas del reloj; debido a su diseño, cualquier pequeña perturbación que cause una oscilación en torno al eje principal 3 se incrementará rápidamente. Las fuerzas de rozamiento que actúan sobre la piedra durante la oscilación detienen el giro y determinan que se inicie otro en sentido opuesto. Una vez que ha comenzado la rotación en sentido inverso, las fuerzas de rozamiento detienen la oscilación. Así pues, el giro en sentido contrario al de las agujas del reloj permite que cierto tipo de inestabilidad –la oscilación en torno al eje principal 3– aumenten en una forma rápida al tiempo que las fuerzas de fricción que acompañan a las oscilaciones inducen la inversión del sentido de giro.



Posibilidades rotatorias con el semihuevo de Moore

Si la piedra está girando inicialmente en sentido horario, se favorece otra inestabilidad, la que genera oscilación alrededor del otro eje horizontal principal. De nuevo, el diseño del rattleback hace que esta inestabilidad crezca exponencialmente en amplitud. Las fuerzas de fricción durante este tipo de oscilación hacen que se detenga el giro en sentido de las agujas del reloj, y se reinicie en sentido contrario.

Quise investigar cómo cambiaría el comportamiento de un rattleback si modificaba una de sus características fundamentales de diseño. La única característica que pude cambiar sin mayores dificultades fue el momento de inercia relativo respecto a los dos ejes horizontales principales. Para cambiarlo sujeté con cinta adhesiva un lápiz sobre la parte superior del rattleback; primero, paralelo a lo largo del rectángulo, y, luego, paralelo a su anchura. En cada prueba tuve cuidado de equilibrar el lápiz de manera que el rattleback girara en la misma zona del elipsoide.

Con el lápiz a lo largo, el rattleback se comportaba, en líneas generales, como lo había hecho anteriormente, puesto que el lapicero apenas aumentaba la diferencia de los momentos de inercia alrededor de los ejes principales. Con el lápiz a lo ancho, la piedra no invirtió su giro, ni mostró alguna de sus inestabilidades características. Situando el lápiz a lo ancho del rectángulo aumentaba el momento de inercia respecto al eje principal 3. Según se desprendía, aumentaba uno lo suficientemente como para hacerlo comparable con el otro. Cuando los momentos de inercia son parecidos, las oscilaciones debidas a pequeñas perturbaciones no crecen exponencialmente con el tiempo, sino que permanecen suaves. En ausencia de grandes oscilaciones alrededor de los ejes principales, las fuerzas de rozamiento con la superficie de la mesa no pueden detener el rattleback ni, en consecuencia, invertir su giro.

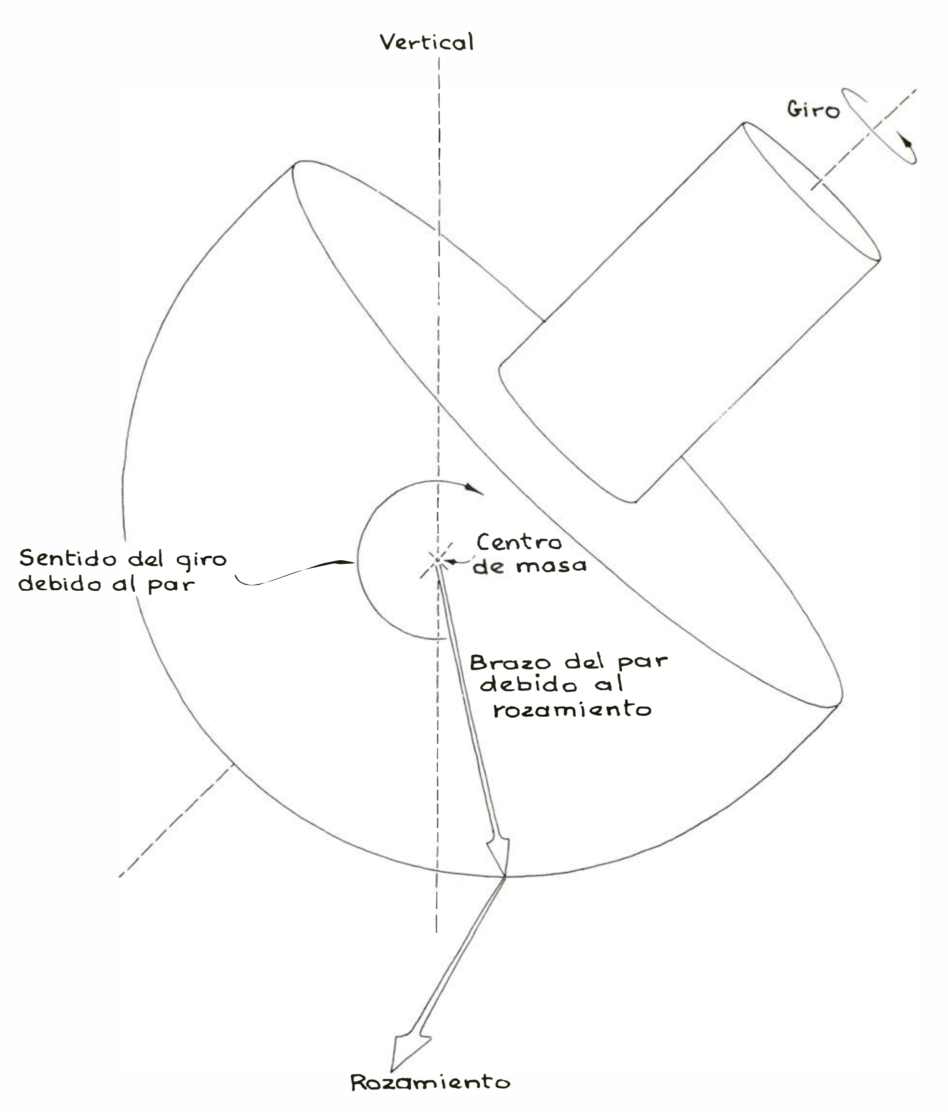
Un momento de inercia puede variarse de una forma más controlada montando pernos largos sobre el rattleback. Se situará un perno en cada sentido a lo largo de un eje principal y se roscan varias tuercas en sus extremos. Se modifica el momento de inercia moviendo las tuercas, aproximándolas al centro del rattleback, o alejándolas de él. (La piedra debe permanecer de todas formas equilibrada.) Con esta disposición conseguí, otra vez, eliminar las inversiones de giro cuando igualaba los momentos de inercia respecto de los dos ejes horizontales principales. ¿Qué suce-

dería si se ajustaran las tuercas para hacer que el momento de inercia alrededor del eje principal 2 fuera mayor que el correspondiente al eje principal 3? ¿Volvería a aparecer la inversión del giro, pero con los sentidos de las inversiones fuerte y débil cambiados?

El rattleback con forma de semihuevo de Moore tiene una varilla de latón por medio de la cual pueden alterarse los momentos de inercia (al cambiar de posición la varilla, se modifica la posición de los ejes principales). Al sujetar la varilla sobre la parte más estrecha de la cara plana, el huevo gira establemente en ambos sentidos. Con la varilla sobre el eje más largo de la cara plana, aparecen inversiones de giro e inestabilidades. Normalmente, se sitúa la varilla formando un pequeño ángulo con el eje más largo de la cara plana, según se muestra en la figura de la página opuesta. El semihuevo presenta una inversión débil en cada sentido de giro. Con la varilla montada de esta forma, el semihuevo es sensible también a golpecitos en su perímetro. El sentido del giro resultante depende del sitio donde se le golpee. La figura indica las distintas posibilidades.

El análisis de G. T. Walker predice que la frecuencia de la oscilación es, alrededor del eje principal 3, mayor que alrededor del eje 2, lo cual es fácilmente comprobable con un rattleback elemental. De sus ecuaciones se desprende, asimismo, que no aumenta ninguna inestabilidad si la frecuencia del giro es mayor que la frecuencia de la oscilación. Con los rattlebacks de Moore, mis esfuerzos por eliminar la inestabilidad en la inversión fuerte (la correspondiente al giro en sentido contrario al de las agujas del reloj) resultaron vanos, ya que mi impulso de giro inicial introducía inevitablemente inestabilidades. Sin embargo, fácilmente podía girar el rattleback en el sentido de las agujas del reloj con una frecuencia de giro mayor que la frecuencia de oscilación alrededor del eje principal 2. Por lo que yo mismo he podido comprobar, la oscilación no aparece hasta que el rozamiento ha menguado la velocidad de giro de la piedra, quizá hasta una frecuencia menor que la frecuencia de oscilación.

He jugado con rattlebacks de Moore sobre superficies pulidas de formica y sobre algunas otras. No he investigado los efectos del coeficiente de rozamiento entre el rattleback y la superficie en la que gira. Si la fuerza de rozamiento es muy alta, el giro desaparecerá tan deprisa que no dará tiempo a que se desarrollen inversiones de giro. Pero si el rozamiento es muy pequeño, la piedra apenas invertirá su giro, toda vez que la inversión precisa de la fricción.



Lo que ocurre cuando se invierte una peonza Tippe

Se pueden tallar rattlebacks en madera o cemento, después se ajustará la forma de la parte inferior hasta que se comporte correctamente. Podemos introducir variaciones en el diseño básico. Vale la pena, por ejemplo, investigar el efecto de la distribución de peso, sujetando con cinta adhesiva pesas o varillas metálicas a la superficie superior. Construyendo varios rattlebacks, lograremos determinar de qué modo afecta a la inversión del giro la orientación del elipsoide. Si se pretende ganar un concurso por el número de inversiones de giro de un rattleback hay que buscar el ángulo óptimo entre el eje mayor del cuerpo y el eje longitudinal del elipsoide. Los más osados podrían construir un rattleback suficientemente grande como para montarlo; si les saliera uno con una inversión de giro vigorosa, montarlo sería como cabalgar sobre un potro salvaje.

Cambiemos ahora al tipo Tippe Top, o Peonza Tippe, que invierte también inesperadamente su giro. Este modelo suele tener la parte inferior esférica y un vástago central. Se la hace bailar como si fuera un trompo corriente sobre una superficie plana. La sorpresa llega cuando la peonza gira sobre su extremo esférico sólo durante unos segundos y, luego, cabecea boca abajo, dando vueltas sobre su vástago. El movimiento parece violar la ley de la conservación de la energía, ya que la peonza levanta su centro de gravedad (que se encuentra en la parte esférica) sin ayuda exterior.

La peonza ha fascinado durante largo tiempo a los observadores, eminentes físicos y matemáticos incluidos. En un reciente trabajo de Richard J. Cohen, del Instituto de Tecnología de Massachusetts, describe cómo William Thomson (genial físico conocido con el sobrenombre de Lord Kelvin) perdía el tiempo haciendo girar piedras pulimentadas sobre la playa, en lugar de preparar sus exá-

menes de matemáticas en la Universidad de Cambridge. Más tarde Niels Bohr, quien desarrolló el primer modelo moderno del átomo de hidrógeno, se sintió igualmente atraído por la mecánica de la peonza de Tippe.

La sorprendente inversión de la peonza se origina debido al rozamiento entre la misma y la superficie sobre la cual gira. Supongamos que gira con su vástago inclinado respecto al eje vertical, como puede apreciarse en la figura de la página anterior. Debido a su giro, aparece una fuerza de rozamiento deslizante (perpendicular al plano de la página en la figura). Esta fuerza origina un par respecto al centro de gravedad, que causa la inversión de la peonza.

Podemos fabricar una peonza de Tippe cortando parte de una pelota de goma maciza e introduciendo un perno en la superficie plana. Un típico aro de niño con una piedra pulida desarrolla el mismo tipo de inversión; basta, para ello, hacer girar el aro en la piedra. Un huevo duro al que se hace girar en posición longitudinal llegará a levantarse para girar sobre uno de sus extremos (en realidad es una de las formas de saber si está bien cocido). Sin embargo, un huevo crudo no se levantaría debido al movimiento de su líquido interno.

En la Feria Internacional de Ciencia e Ingeniería celebrada este año en San Antonio, Texas, hubo varios proyectos que trataban temas discutidos recientemente en esta sección. En la sección de física, David Kirk, de Edmond, trató de determinar si el agua inicialmente caliente podría, bajo determinadas circunstancias, congelarse mucho antes que la inicialmente templada. Sus resultados fueron, como los míos, muy variados, pero éstos parecían indicar que con recipientes destapados se puede demostrar el efecto. Joella Carr, de Winchester, estudió la flotación de gotas de agua sobre la superficie de una masa de agua, confirmando que la vibración de dicha masa prolonga la duración de las gotas durante diez minutos más o menos.

En la sección de química se ofrecieron investigaciones sobre reacciones químicas oscilantes. Helen Rumels, de Baton Rouge, estudió la periodicidad de la reacción de Belousov, llegando a la conclusión de que aumentando la concentración de cerio de 0 a 10^{-4} molar se podía incrementar el período de la triple reacción. También expuso que las contaminaciones de cloruro, bromuro o yoduro detenían la oscilación. Su indicador era nitroferroín. William Crooke, de Pensacola, investigó sobre las ondas químicas en una reacción similar. Science Service, 1719 N. Street, Washington, D. C. 20036, vende los proyectos.

Libros

La especie y la especiación en el reino animal y un atlas gráfico de ultraestructura celular

Antonio Prevosti y Mercedes Durfort

LES PROBLÈMES DE L'ESPÈCE DANS LE REGNE ANIMAL, por Ch. Bocquet, J. Génermont y M. Lamotte, vol. I. Memoires n.º 38 de la Société Zoologique de France, 408 págs. París 1976. Vol. II. Memoires n.º 39 de la Société Zoologique de France, 384 págs. 1977. El capítulo I, a cargo de los directores de la obra, es una introducción general al problema de la especie en el reino animal. Después de una rápida reseña histórica sobre el concepto de especie (esta reseña presenta un fuerte sesgo por recoger casi exclusivamente las ideas de autores franceses), plantean el problema en los términos actuales. Dentro de una población existen diferencias individuales (polimorfismo); las poblaciones de una misma especie pueden presentar caracteres diferenciales (politipismo). Estas situaciones, así como el polimorfismo sexual, pueden tener por consecuencia que, si faltan observaciones adecuadas en la naturaleza, lleguen a considerarse de diferente especie individuos ilimitadamente interfecundos. Una complicación opuesta es que individuos pertenecientes a grupos de reproducción separados no se distinguan o sean muy difíciles de distinguir por su aspecto, formando especies gemelas. Consideran que el proceso de especiación es una consecuencia de la progresiva diferenciación genética en las poblaciones de una especie, que pueden convertirse en razas geográficas y finalmente en especies diferentes si se produce aislamiento reproductor.

El proceso de esta progresiva diferenciación lo explican en función de los mecanismos de la genética de poblaciones. En relación con ello, plantean la cuestión de que las mutaciones génicas no afectan, a nivel citológico, a las posibilidades de hibridación entre los individuos y, por otra parte, que la diferenciación morfológica no es siempre paralela a la divergencia genética, pues ésta puede incidir principalmente en caracteres fisiológicos o bioquímicos, no accesibles a la observación directa. Esto tiene por consecuencia que los grados de diferenciación morfológica, fisiológica y citológica entre distintas especies pueden ser muy diferentes y no estar correlacionados.

Aunque estiman que el criterio de interfecundidad es el más importante para diagnosticar que dos individuos o poblaciones pertenecen a la misma especie, analizan también las dificultades en la utilización de este criterio. Los círculos de razas son un ejemplo de dichas dificultades, pero lo más frecuente es que las poblaciones estén separadas y sus individuos no se encuentren en la naturaleza para probar su interfecundidad. Como es sabido, en este caso las pruebas de laboratorio no resuelven el problema, ya que las condiciones artificiales pueden alterar lo que ocurría espontáneamente en la naturaleza.

Partiendo de estos planteamientos, distintos autores estudian la situación en el grupo animal en que trabajan, atendiendo a las peculiaridades de dicho grupo y a los criterios impresos por la personalidad de los especialistas destacados en la elaboración de su sistemática. Por el sentido realista con que se estudia el estado actual de la sistemática de los grupos zoológicos, esta obra es única y original.

F. Vuilleumier considera, en el segundo capítulo del primer tomo, la noción de especie en ornitología. Analiza estadísticamente las descripciones de especies nuevas entre 1955 y 1974, así como los casos concretos de algunos géneros y subgéneros. Concluye que en las Aves la noción de especie biológica sigue siendo "teórica". Ya que la especie "práctica" continúa fundándose en la morfología, en la tipológica. Discute también el hecho de que, por no ser posible estimar la diferenciación genética a base de la morfológica, los ornitólogos han creado una serie de categorías como los Rassenkreis, semiespecies, Formenkreiss, aloespecies, etcétera.

La especie en el ámbito de los Teleósteos es estudiada por J. Daget y M. L. Bauchot. Plantean primero el problema desde el punto de vista de la necesidad de conocer la variabilidad morfológica existente dentro de las especies, concretamente los estados ontogenéticos (relaciones de alometría de crecimiento, características de las formas juveniles, desarrollo con metamorfosis), las variaciones hereditarias (de color, morfológicas, merísticas, patológicas etcétera) y el dimorfismo sexual. Después atienden a los factores biológicos que condicionan la especiación: poliploidía, cambios estructurales en los cromosomas, aislamiento fisiológico y etológico, aislamiento geográfico y ecológico. Terminan el trabajo discutiendo los criterios y técnicas utilizadas en la sistemática del grupo, proyecciones, dendrogramas, cladogramas y taxonomía numérica.

Guillaumin y Descimon tratan de la especie en los Lepidópteros. Comparan a este grupo, a veces hasta en la cantidad de aficionados que se ocupan de él, al de las Aves, analizando los criterios principales que se utilizan en su sistemática y discutiendo el peso real que suele dárseles. Dichos criterios son el hábito (principalmente dibujos y coloración de las alas), los genitalia, el cariotipo, los caracteres morfológicos de larvas y ninfas, los caracteres biométricos, la investigación de la existencia de híbridos en la naturaleza y los cruzamientos en el laboratorio. Como ejemplo de la utilización de estos criterios analizan algunos casos concretos.

Los Drosofílidos ocupan un lugar destacado en los estudios de sistemática moderna, ya que el concepto moderno de especie gonocórica nació de la comprensión mutua entre sistemáticos que asimilaron los descubrimientos de la genética y genéticos a los que el trabajo de laboratorio no les hizo olvidar la existencia de poblaciones naturales. Esta es la opinión de Tsacas y Boquet, quienes abordan el estudio del grupo. Describen los principales caracteres morfológicos utilizados en la sistemática del mismo, así como la utilización del cariotipo, cromosomas gigantes y caracteres etológicos. Terminan considerando, principalmente con el ejemplo de las "razas" de *Drosophila paulistorum*, el problema de las especies "in statu nascendi".

La sistemática de los mosquitos ha evolucionado mucho en los últimos 30 años, debido a una estrecha colaboración entre sistemáticos clásicos, ecólogos y genéticos. Se ha visto que muchas de las antiguas especies son, en realidad, un complejo de especies gemelas o de subespecies alopátricas que pueden distinguirse más o menos fácilmente. Por ello, los sistemáticos de este grupo ya no pueden trabajar con material de museo, debiendo hacerlo en el campo y en el labo-

ratorio, realizando estudios morfológicos, ecológicos, etológicos, genéticos y biogeográficos. Grjebine, Coz, Eluard, Mouchet y Rageau, presentan los resultados de esta nueva orientación de la sistemática del grupo, estudiando los casos de los grupos de *Aedes (Stegomya) scutellaris, Anopheles maculipennis, Anopheles gambiae* y *Culex pipiens*.

Según Bocquet, también en los Crustáceos el criterio de hibridación debe tomarse como el único decisivo, pero desgraciadamente pocas veces su utilización es posible. Entre estos pocos casos están los géneros *Tisbe* (Copépodos) y el complejo de *Jaera albifrons* (Isópodos), en los que se apoya dicho autor para explicar de qué modo la sistemática evolutiva enfoca, a propósito de este grupo, el problema de la especie.

Los Lamelibranquios forman un grupo muy homogéneo, lo cual unido a la gran variabilidad morfológica y de color que presentan algunas especies, a menudo hace difícil su distinción. Casi toda la sistemática del grupo se basa en caracteres de la concha, debido a la facilidad con que ésta se conserva; en cambio, los trabajos basados en la ecología, la hibridación, el desarrollo y otros caracteres biológicos son escasos. P. Lubet, autor del capítulo que trata de estos moluscos, estudia primero la morfología de la concha del adulto, pasando luego a los caracteres fisiológicos, bioquímicos y cariotípicos y a la hibridación.

El último capítulo del primer tomo estudia la especie en los protozoos. Debido a su pequeño tamaño, se hace difícil apreciar las diferencias morfológicas dentro de este grupo. No obstante, hay algunos casos descritos con detalle. El autor del capítulo, J. Génermont, ha elegido para presentar los problemas de la especie en los protozoos el "complejo de *Paramaecium aurelia* y el de *Euplotes vannus* como ejemplos de formas con reproducción sexual, en las cuales el estudio de los sistemas de incompatibilidad sexual y la interfecundidad ha puesto de manifiesto la insuficiencia del criterio tipológico de especie. Dedica la última parte de su trabajo a las formas sin reproducción sexual, principalmente a las amibas del género *Naegleria*. Concluye que es ilusorio definir las especies de este género desde un punto de vista biológico, mientras que es operativo dar un nombre a las distintas formas patógenas de acuerdo con su importancia epidemiológica.

El segundo tomo comienza con un trabajo de Blandin sobre las arañas. Afirma que la sistemática de este grupo está todavía en la fase de inventario de especies morfológicas. Analiza primero los datos sobre descripción de especies nuevas desde 1758 a 1969 y luego pasa a detallar los caracteres que se utilizan en el diagnóstico: forma y ornamentación, ojos, espinas y genitalia. También considera los datos biogeográficos, estudiando varios ejemplos; y recurre al criterio mixiológico tanto en la comparación de especies alopátricas como en relación con las barreras entre especies simpátricas.

De los Colémbolos se ocupa Cassagnau, quien se extiende especialmente en la importancia de los cromosomas politénicos, aunque se refiere también a los criterios morfológicos, señalando la importancia de los caracteres no adaptativos, como la quetotaxia, para establecer relaciones genealógicas. Señala igualmente las dificultades que se originan por culpa de la maleabilidad de los caracteres morfológicos, debido al crecimiento diferencial, la epitoquia (diferencias morfogenéticas relacionadas con la fase de reproducción), las ecomorfosis, las ciclomorfosis, etcétera.

La sistemática de los Ortópteros está actualmente muy trabajada, habiendo penetrado en ella bastante los conceptos de la "nueva sistemática". Ph. Dreux inicia su colaboración afrontando el polimorfismo de muchas especies que, particularmente cuando se manifiesta en fases distintas, plantea dificultades al sistemático. Luego analiza los complejos de especies gemelas que se han ido descubriendo con la utilización de los criterios de la "nueva sistemática"; hace notable hincapié en los mecanismos de aislamiento entre especies. Después de pasar revista a los criterios cariológicos, termina con una exposición de la teoría de la "especiación estasipátrica" propuesta por White, precisamente basándose en datos de estos insectos.

Los Gasterópodos presentan características parecidas a los de los Lamelibraquios, en lo que se refiere a la importancia de la concha. Como carácter morfológico también se utiliza mucho en ellos la rádula. A diferencia de los Lamelibranquios, especialmente en las formas terrestres y dulceacuícolas, en este grupo se han esgrimido con frecuencia los caracteres genéticos: cariotipo e hibridación. A. Franc explica cómo se han utilizado estos caracteres, junto con los de reproducción y desarrollo, la especificidad parasitaria y los caracteres ecológicos.

A. Dubois analiza los Anuros, grupo al que han prestado mucha atención los estudiosos de la evolución. En un extenso capítulo de más de 100 páginas, tras exponer algunos complejos clásicos como el complejo de *Rana pipiens* y el de *Rana esculenta*, atendiendo a los distintos criterios que se han empleado para su estudio, pasa a considerar los caracteres de importancia general para la sistemática del grupo. Así considera la poliploidía, la morfología, la biometría y el color de los adultos, las fases del desarrollo, los caracteres bioquímicos, el canto, la ecología y la etología, la hibridación y la distribución en el espacio y el tiempo. Termina con una amplia discusión del concepto biológico de especie y de la nomenclatura zoológica.

Los Urodelos también han sido bastante estudiados, aunque están mucho menos diversificados que los anuros. Gasser, que analiza este grupo, trata de la variación genética intraespecífica (polimorfismos enzimáticos y cromosómicos), la variación geográfica, los mecanismos de aislamiento y la hibridación. Aborda, por fin, los factores que dirigen la especiación en el grupo.

G. Pasteur comenta el problema de la especie de los Lygodactylos, centrándose en el género *Lygodactylus*. A pesar de no haberse conseguido experimentos de cruce con las distintas formas existentes dentro de la antigua especie *Lygodactylus picturatus*, considera que ésta puede compararse a *Drosophila paulistorum*, formando una superespecie, que presenta grados y modos de diferenciación sistemática muy distintos. A falta de los experimentos de cruce, los principales criterios utilizados para distinguir estos grupos son los caracteres de pigmentación y los cortejos nupciales.

En el último capítulo del segundo tomo, J. Chaline y L. Thaler analizan el problema de la especie en los roedores desde los puntos de vista biológico y paleontológico. El enfoque es muy general. Tratan de la distinción de especies simpátricas, atendiendo a la citogenética y al estudio de la variabilidad aloenzimática, de las especies parapátricas y de la especie en el tiempo.

Aunque esta obra interesa especialmente a los zoólogos, también los profesionales de la genética de poblaciones y de la evolución en general encontrarán recopilados en ella una gran cantidad de datos, descripción de situaciones, puntos de vista, criterios e interpretaciones que se presentan juntos por primera vez. Esta gran variedad de información, junto con el punto de vista ecléctico con que se trata, dan una visión muy amplia del problema de la especiación y de la utilización del concepto de especie en la práctica de la sistemática. La originalidad de los dos tomos publicados permite

esperar con interés la aparición del tercero, que debe concluir la obra, en el que se sintetizará lo descrito en los dos primeros. (A.P.)

ULTRAESTRUCTURA DE LAS CELULAS DE LOS MAMIFEROS, por Radivoj V. Krstić. Ed. Eunibar, Barcelona 1979. En la interpretación de cualquier preparación, además de los conocimientos sobre los tejidos y las células que la componen, debe tenerse presente la metodología seguida hasta la obtención de lo que se observa. Cuando el estudio del material biológico se lleva a cabo con el microscopio electrónico, contemporáneamente a los factores clásicos hemos de prestar atención a los posibles artefactos introducidos en el material tras la fijación e inclusión, así como a los aportados con el contrastado de los cortes ultrafinos, ya que dada la elevada resolución del instrumento (de 1 a 10 angstrom) se detecta cualquier alteración por pequeña que sea.

Para la interpretación debe tenerse muy presente la orientación del corte observado. Efectivamente, las secciones obtenidas con el ultramicrótomo son del orden de los 300 a 400 angstrom de grosor; ello indica que una estructura esférica u ovoidea, una mitocondria por ejemplo, podrá ser cortada en numerosos planos atendiendo a la orientación del corte efectuado, y según como se hallaba el orgánulo en la célula se obtendrán distintas imágenes no siempre fáciles de reconocer, a no ser por la doble membrana que presenta el orgánulo en cuestión; en el caso de estructuras laberínticas, como es el aparato de Golgi, las imágenes obtenidas al microscopio electrónico son múltiples.

En definitiva, aparte de los conocimientos teóricos y técnicos que ha de poseer el estudioso de estos temas, para interpretar cualquier imagen al microscopio electrónico, debe tener intuición y estar dotado de imaginación para conjugar simultáneamente todos los factores. Evidentemente, la experiencia facilita enormemente esta labor.

De ahí se desprende el elevado interés de *Ultraestructura de las células de los mamíferos*, para ayudar a ejercitar la imaginación, al mostrar cómo se disponen las estructuras citoplasmáticas en su contexto, es decir: la tridimensionalidad que no ofrece, ni puede ofrecernos, el microscopio electrónico, instrumento que por sus características no permite la disección de la imagen. El autor lo consigue mediante unos esquemas muy bien logrados, y sobre todo, salvo algunas excepciones, muy proporcionados, detalle este de gran importancia ya que en obras muy prestigiosas se han deslizado esquemas en los que el tamaño de un centriolo puede llegar a triplicar el de una mitocondria cortada transversalmente por la mitad.

La obra de Krstić, profesor de la Universidad de Lausana, comprende 176 láminas de esquemas de las diversas estructuras celulares observadas al microscopio electrónico, desde la célula epitelial, en sus distintas modalidades, a la neurona, tratando asímismo de células altamente especializadas como son el espermatozoide y las células fotosensibles de la retina. Cada lámina va acompañada de un texto explicativo, cuya traducción ha sido efectuada con corrección, salvo algún término, que no deja de sorprendernos, a pesar de leerse con cierta regularidad y estar acostumbrados a una determinada anarquía terminológica existente en el campo de la ultraestructura celular, al igual que en otras materias, siendo generalmente las traducciones las responsables de la adopción de vocablos no demasiado adecuados. Así, en este texto, puede leerse "organitos" en lugar de *orgánulos*, "teñidos" en lugar de *contrastados* (nombre correcto de la técnica habitualmente seguida tras la obtención de los cortes ultrafinos y previa observación de los mismos), "campos de Golgi" en lugar de *dictiosomas*, "microvellosidades" en lugar de *microvilli*, "láminas anulares" en lugar de *láminas anilladas* o *periódicas*.

Es interesante destacar, entre otras, las imágenes de los esquemas tridimensionales del aparato de Golgi (páginas 63 y 67), del ergastoplasma (páginas 71, 95 y 99), así como el de las láminas anilladas (página 111), centriolo (página 115), etcétera. Lo ideal hubiera sido que la obra estuviera dotada de una selección de microelectrografías, si bien existen distintas ediciones de atlas de ultraestructura celular que en un 95 por ciento están ilustradas con profusión de imágenes y que en cambio sólo aportan escasos esquemas. Posiblemente, la obra de Krstić es la única que, con estructura típica de atlas, está formada en su totalidad por esquemas.

El libro en cuestión constituye un auxiliar didáctico de primera categoría para la interpretación de los cortes ultrafinos por parte de estudiantes de biología, farmacia, medicina y veterinaria, así como para los más conspicuos expertos en estudios ultraestructurales.

El hecho de que el prólogo, aunque breve, sea del profesor W. Bargmann, catedrático de anatomía de la Universidad de Kiel y eminente histólogo, supone una garantía del valor de la obra. (M.D.)

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

DIFERENCIAS DE COLOR

COLOR IN BUSINESS, SCIENCE AND INDUSTRY. Deane B. Judd y Günther Wyszecki. Segunda edición. J. Wiley & Sons, N. York, 1963.

PRINCIPLES OF COLOR TECHNOLOGY. Fred W. Billmeyer Jr. y Max Saltzman. J. Wiley & Sons, N. York, 1966.

COLOR SCIENCE. Günther Wyszecki y W. S. Stiles, J. Wiley & Sons, N. York, 1967.

EL COLOR Y SU MEDICION. Roberto Daniel Lozano. Americalee, Buenos Aires, 1978.

CONDUCTORES DE CADENA LINEAL

SYNTHESIS AND PROPERTIES OF LOWDIMENSIONAL MATERIALS. Dirigido por Joel S. Miller y Arthur J. Epstein en *Annals of the New York Academy of Sciences,* vol. 313; 29 de septiembre de 1978.

QUASI-ONE-DIMENSIONAL CONDUCTORS. Dirigido por S. Barišič, A. Bjeliś, J. R. Cooper y B. Leontič en *Lecture Notes in Physics,* vols. 95 y 96. Springer-Verlag, 1979.

EL MICROSCOPIO ACUSTICO

ACOUSTIC MICROSCOPY: BIOMEDICAL APPLICATIONS. Ross A. Lemons y Calvin F. Quate en *Science,* vol. 188, n.º 4191, págs. 905-911; 30 de mayo de 1975.

ACOUSTIC IMAGING: CAMERAS, MICROSCOPES, PHASED ARRAYS, AND HOLOGRAPHIC SYSTEMS. Dirigido por Glen Wade. Plenum Press, 1976.

ACOUSTIC MICROSCOPY AT OPTICAL WAVELENGTHS. V. Jipson y C. F. Quate en *Applied Physics Letters,* vol. 32, n.º 12, págs. 789-791; 15 de junio de 1978.

AN ANGULAR-SPECTRUM APPROACH TO CONTRAST IN REFLECTION ACOUSTIC MICROSCOPY. Abdullah Atalar en *Journal of Applied Physics,* vol. 49, n.º 10, págs. 5130-5139; oct. 1978.

LEGIONELOSIS

LEGIONNAIRES' DISEASE: DESCRIPTION OF AN EPIDEMIC OF PNEUMONIA. David W. Fraser, Theodore F. Tsai, Walter Orenstein, William E. Parkin, H. James Beecham, Robert G. Sharrar, John Harris, George F. Mallison, Stanley M. Martin, Joseph E. McDade, Charles Shepard, Philip S. Brachman y el equipo de investigación de campo en *The New England Journal of Medicine,* vol. 297, n.º 22, págs. 1189-1197; 1 de diciembre de 1977.

LEGIONNAIRES' DISEASE: ISOLATION OF A BACTERIUM AND DEMONSTRATION OF ITS ROLE IN OTHER RESPIRATORY DISEASE. Joseph E. McDade, Charles Shepard, David W. Fraser, Theodore F. Tsai, Martha A. Redus, Walter R. Dowdle y el equipo de investigación de laboratorio en *The New England Journal of Medicine,* vol. 297, n.º 22, págs. 1197-1203; 1 de diciembre de 1977.

PONTIAC FEVER: AN EPIDEMIC OF UNKNOWN ETIOLOGY IN A HEALTH DEPARTMENT. I. CLINICAL AND EPIDEMIOLOGIC ASPECTS. Thomas H. Glick, Michael B. Gregg, Bernard Berman, George Mallison, Wallace W. Rhodes, Jr., e Ira Kassanoff en *American Journal of Epidemiology,* vol. 107, n.º 2, págs. 149-160; febrero, 1978.

LA MEMBRANA FOTOSINTETICA

A PARTICLE SPANNING THE PHOTOSYNTHETIC MEMBRANE. Kenneth R. Miller en *Journal of Ultrastructure Research,* vol. 54, n.º 1, págs. 159-167; enero, 1976.

DYNAMIC STRUCTURAL FEATURES OF CHLOROPLAST LAMELLAE. Charles J. Arntzen en *Current Topics in Bioenergetics,* vol. 8, págs. 111-160; 1978.

LOS LAGOS DE LAVA DEL KILAUEA

PETROLOGY OF THE KILAUEA IKI LAVA LAKE, HAWAII. Donald H. Richter y James G. Moore. Geological Survey Professional Paper 537-B, 1966.

THE ERUPTION OF AUGUST 1963 AND THE FORMATION OF ALAE LAVA LAKE, HAWAII. Dallas L. Peck y W. T. Kinoshita. Geological Survey Professional Paper 935-A, 1976.

EVOLUCION DE LOS CROCODILIOS

THE CROCODILIAN SKULL IN HISTORICAL PERSPECTIVE. Wann Langston, Jr., en *Biology of the Reptilia:* vol. 4, dirigido por Carl Gans y Thomas S. Parsons. Academic Press, 1973.

CROCODYLIA. Rodney Steel en *Handbuch der Paläoherpetologie/Encyclopedia of Paleoherpetology:* parte 16, dirigido por Oskar Kohn. Gustav Fischer Verlag, 1973.

THE MORPHOLOGY AND RELATIONSHIPS OF A CROCODILIAN, *ORTHOSUCHUS STORMBERGI,* FROM THE UPPER TRIASSIC OF LESOTHO. Diane S. Nash en *Annals of the South African Museum,* vol. 67, parte 7, págs. 227-329; 1975.

MATEMATICA CONSTRUCTIVISTA

THE FOUNDATIONS OF CONSTRUCTIVE ANALYSIS. Errett Bishop. McGraw-Hill Book Company, 1967.

WHAT IS A REAL NUMBER? John Myhill en *The American Mathematical Monthly,* vol. 79, n.º 7, págs. 748-754; agosto-septiembre, 1972.

A RENAISSANCE OF EMPIRICISM IN THE RECENT PHILOSOPHY OF MATHEMATICS. Imre Lakatos en *Mathematics, Science and Epistemology.* Cambridge University Press, 1978.

JUEGOS MATEMATICOS

ON PACKING SQUARES WITH EQUAL SQUARES. P. Erdös y R. L. Graham en *Journal of Combinatorial Theory, Series A,* vol. 19, n.º 1, págs. 119-123; julio, 1975.

INEFFICIENCY IN PACKING SQUARES WITH UNIT SQUARES. L. F. Roth y R. C. Vaughan en *Journal of Combinatorial Theory, Series A,* vol. 24, n.º 2, págs. 170-186; marzo, 1978.

TALLER Y LABORATORIO

AN ELEMENTARY TREATMENT OF THE THEORY OF SPINNING TOPS AND GYROSCOPIC MOTION. Harold Crabtree. Longmans, Green and Co., 1909.

A TREATISE ON GYROSTATICS AND ROTATIONAL MOTION: THEORY AND APPLICATION. Andrew Gray. Dover Publications, Inc., 1959.

MECHANICS: LECTURES OF THEORETICAL PHYSICS, vol. 1, Arnold Sommerfeld. Academic Press, 1964.

THE TIPPE TOP REVISITED. Richard J. Cohen en *American Journal of Physics,* vol. 45, n.º 1, págs. 12-17; enero, 1977.

Indice anual

Este índice contiene los nombres de los autores y artículos aparecidos en Investigación y Ciencia *a partir de enero de 1979. Se incorpora asimismo la relación alfabética de las secciones* Juegos matemáticos *y* Taller y laboratorio.

AUTORES

ARTICULOS

TALLER Y LABORATORIO

Autor: Jearl Walker

JUEGOS MATEMATICOS

Autor: Martin Gardner